



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 123946

TO: Phillip Gambel
Location: 3e81 / 3c70
Wednesday, June 09, 2004
Art Unit: 1644
Phone: 272-0844
Serial Number: 09 / 891943

From: Jan Delaval
Location: Biotech-Chem Library
Rem 1A51
Phone: 272-2504

jan.delaval@uspto.gov

Search Notes

Delaval, Jan

123946

From: Gambel, Phillip
Sent: Monday, June 07, 2004 9:23 AM
To: Delaval, Jan
Subject: 09/891,943 sequence search

jan

please perform a sequence and a sequence interference search for

09 / 891,943 (gallatin)

SEQ ID NO: 1

SEQ ID NO: 2

thanx

phillip gambel
art unit 1644
272-0844

1644 mailbox 3c70

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: _____ Examiner #: _____ Date: _____
 Art Unit: _____ Phone Number 30 _____ Serial Number: _____
 Mail Box and Bldg Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or novelty of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc. if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

Earliest Priority Filing Date: _____

For Sequence Searches Only Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

STAFF USE ONLY

	Type of Search	Vendors and cost where applicable
Searcher: <u>per</u>	NA Sequence (#) <input checked="" type="checkbox"/>	STN _____
Searcher Phone #: <u>22504</u>	AA Sequence (#) <input checked="" type="checkbox"/>	Dialog _____
Searcher Location: _____	Structure (#) _____	Questel/Orbit _____
Date Searcher Picked Up: <u>6/7</u>	Bibliographic _____	Dr. Link _____
Date Completed: <u>6/9</u>	Litigation _____	Lexis/Nexis _____
Searcher Prep. & Review Time: _____	Fulltext _____	Sequence Systems <input checked="" type="checkbox"/>
Clerical Prep Time: <u>10</u>	Patent Family _____	WWW/Internet _____
Online Fee: <u>+ 20</u>	Other _____	Other (specify) _____

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: June 8, 2004, 05:52:06 ; Search time 22742 Seconds
(without alignments)
7101.229 Million cell updates/sec
Title: US-09-891-943-1
Perfect score: 3726
Sequence: 1 tgaccttcgacactgtcgtt.....agcataaatttcattatgct 3726
Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0
Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
Database : GenEmbl:
1: gb_ba:*
2: gb_atg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3726	100.0	3726	6	AR044695
2	3726	100.0	3726	6	AR052342 Sequence
3	3726	100.0	3726	6	AR055140 Sequence
4	3726	100.0	3726	6	AR158011 Sequence
5	3726	100.0	3726	6	I13528 Sequence 1
6	3726	100.0	3726	6	I15787 Sequence 1
7	3726	100.0	3726	6	I2491 Sequence 1
8	3726	100.0	3726	6	AR222661 Sequence
9	3726	100.0	3726	6	AR399748 Sequence
10	3726	100.0	3726	6	AX448965 Sequence
11	3726	100.0	3726	6	BD070942 Novel hum
12	3699	99.3	3785	6	AR044775 Sequence
13	3699	99.3	3785	6	AR052422 Sequence
14	3699	99.3	3785	6	AR055220 Sequence
15	3699	99.3	3785	6	AR158091 Sequence
16	3699	99.3	3785	6	I2571 Sequence 98
17	3699	99.3	3785	6	AR222741 Sequence
18	3699	99.3	3785	6	AR399828 Sequence
19	3699	99.3	3785	6	AX449062 Sequence
20	3699	99.3	3785	6	BD071022 Novel hum
21	3699	99.3	3785	6	AR044774 Sequence
22	3583	96.2	3956	6	AR052421 Sequence
23	3583	96.2	3956	6	AR055219 Sequence
24	3583	96.2	3956	6	AR158090 Sequence
25	3583	96.2	3956	6	I92570 Sequence 97
26	3583	96.2	3956	6	AR222740 Sequence
27	3583	96.2	3956	6	AR399827 Sequence
28	3583	96.2	3956	6	AX449061 Sequence
29	3583	96.2	3956	6	BD071021 Novel hum
30	3484.6	93.5	3486	9	HSU37028 U37028 Human leuko
31	2315.2	62.1	3803	6	AR044732 Sequence
32	2315.2	62.1	3803	6	AR052379 Sequence
33	2315.2	62.1	3803	6	AR055177 Sequence
34	2315.2	62.1	3803	6	AR158048 Sequence
35	2315.2	62.1	3803	6	I92528 Sequence 52
36	2315.2	62.1	3803	6	AR222698 Sequence
37	2315.2	62.1	3803	6	AR399785 Sequence
38	2315.2	62.1	3803	6	AX449016 Sequence
39	2315.2	62.1	3803	6	BD070979 Novel hum
40	2304.2	61.9	3558	10	AF021334 Rattus no
41	2304.2	61.8	3597	6	AR044733 Sequence
42	2304.2	61.8	3597	6	AR052380 Sequence
43	2304.2	61.8	3597	6	AR055178 Sequence
44	2304.2	61.8	3597	6	AR158049 Sequence
45	2304.2	61.8	3597	6	I92529 Sequence 54

ALIGNMENTS

RESULT 1
AR044695
LOCUS AR044695 3726 bp DNA linear PAT 29-SRP-1999
DEFINITION Sequence 1 from patent US 5817515.
ACCESSION AR044695
VERSION AR044695.1 GI:5966160
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 3726)
AUTHORS Gallatin,W,Michael, and Van der Vieren,M.
TITLE Human B2 integrin alpha subunit antibodies
JOURNAL Patent: US 5817515-A 1 06-0CT-1998;
FEATURES Location/Qualifiers

source	1. .3726	100.0%; Score 3726; DB 6; Length 3726;	0;	0;	0;
	/organism="unknown"				
	/mol_type="unassigned DNA"				
ORIGIN					
Query Match					
Best Local Similarity 100.0%; Pred. No. 0;					
Matches 3726; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
Qy	1	TGACCTTCGGCAGCTGCTCTTCTGAGTGTCTGCTCTTATCATGATTAACCTGG	60		
Db	1	TGACCTTCGGCAGCTGCTCTTCTGAGTGTCTGCTCTTATCATGATTAACCTGG	60		
Qy	61	ATGTGAGAGGCTACGATCTTCCAGAGAGATGCAGGCGCTTTGGGAGAGCGTGTGC	120		
Db	61	ATGTGAGAGGCTACGATCTTCCAGAGAGATGCAGGCGCTTTGGGAGAGCGTGTGC	120		
Qy	121	AGTTTCGGTGTGATCTCGACTCGTGGTGGGAGCACCTCTGGAGGTGGTGGCGGCAACCCAGA	180		
Db	121	AGTTTCGGTGTGATCTCGACTCGTGGTGGGAGCACCTCTGGAGGTGGTGGCGGCAACCCAGA	180		
Qy	181	CGGAGCGGCTGTATGATCTGCGAGCTGCCACCGGAGATGTCCAGCCCATCCCGCTGCACA	240		
Db	181	CGGAGCGGCTGTATGATCTGCGAGCTGCCACCGGAGATGTCCAGCCCATCCCGCTGCACA	240		
Qy	241	TCCGCTCAGGCGCTGAACATCTCTTGGGCTGACCTCTGGCAGCTCCACCAACGGCT	300		
Db	241	TCCGCTCAGGCGCTGAACATCTCTTGGGCTGACCTCTGGCAGCTCCACCAACGGCT	300		
Qy	301	CCGGCTCTCGGCTGTGGCCGACCTCTGCAAGATCTGTGGGAGAACTATATCTAA	360		
Db	301	CCGGCTCTCGGCTGTGGCCGACCTCTGCAAGATCTGTGGGAGAACTATATCTAA	360		
Qy	361	AGGTTTCCTGCTCTGCTGGCTCGCTGGGAGATCATCCAGACAGTCCCGAGCCCA	420		
Db	361	AGGTTTCCTGCTCTGCTGGCTCGCTGGGAGATCATCCAGACAGTCCCGAGCCCA	420		
Qy	421	CGCCAGAGTCTCCACATCAAGAGATGACATCTCTTCTGATTTGACGCTCTGGAGCA	480		
Db	421	CGCCAGAGTCTCCACATCAAGAGATGACATCTCTTCTGATTTGACGCTCTGGAGCA	480		
Qy	481	TTGACCAAAATGACTTTAAACAGATGAGGCTTTGTCCAAGCTGTCAATGGGCGAGTTG	540		
Db	481	TTGACCAAAATGACTTTAAACAGATGAGGCTTTGTCCAAGCTGTCAATGGGCGAGTTG	540		
Qy	541	AGGGCACTGACACCTCTGTTGCACTGATGAGTACTCAAACTCTCTGAAAGATCCACTCA	600		
Db	541	AGGGCACTGACACCTCTGTTGCACTGATGAGTACTCAAACTCTCTGAAAGATCCACTCA	600		
Qy	601	CTTTCACCAATTCGGACACCGGCTCGGACGAGAGCTGTGGATCCCTCTGCTCCAC	660		
Db	601	CTTTCACCAATTCGGACACCGGCTCGGACGAGAGCTGTGGATCCCTCTGCTCCAC	660		
Qy	661	TGAAAGGCTGAGCTTCAACGGCCACCGGCTATCTGACAGTGTGACACAGCTATTTTCATC	720		
Db	661	TGAAAGGCTGAGCTTCAACGGCCACCGGCTATCTGACAGTGTGACACAGCTATTTTCATC	720		
Qy	721	ATAAGATGGGCGGCGGAAAGTCCAGAGATCTCTATGTCATCAGATGGGCGAGA	780		
Db	721	ATAAGATGGGCGGCGGAAAGTCCAGAGATCTCTATGTCATCAGATGGGCGAGA	780		
Qy	781	AGTACAAAGACCCCTCGAATACAGTGTCTATCCCGGAGAGAGGCTGGGATCA	840		
Db	781	AGTACAAAGACCCCTCGAATACAGTGTCTATCCCGGAGAGAGGCTGGGATCA	840		
Qy	841	TCGCTACGCTATCGGGGTGGGACACGCTTTCAGGGACCCATGCGAGCAGGAGCTGA	900		
Db	841	TCGCTACGCTATCGGGGTGGGACACGCTTTCAGGGACCCATGCGAGCAGGAGCTGA	900		
Qy	901	ATACCATCAGCTCAGCGCTCCGAGGACACGCTTTCAGAGTGGGCAACTTTGACGCGC	960		
Db	901	ATACCATCAGCTCAGCGCTCCGAGGACACGCTTTCAGAGTGGGCAACTTTGACGCGC	960		

Qy	961	TTGGCAGCATCCAGAAAGAGCTGCAGAGAGATCTATGAGTTGAGGAAACCCAGTCCA	1020		
Db	961	TTGGCAGCATCCAGAAAGAGCTGCAGAGAGATCTATGAGTTGAGGAAACCCAGTCCA	1020		
Qy	1021	GGGCAAGCAGCTCTTCCAGCAGAGATGTCCCAAGAGCTTTCAGCAGAGCTTCAAAA	1080		
Db	1021	GGGCAAGCAGCTCTTCCAGCAGAGATGTCCCAAGAGCTTTCAGCAGAGCTTCAAAA	1080		
Qy	1081	TGAGTGGCTCTTCTGGGGCTGTGGGAGCTTTAGCTGTGGTGTGAGGTGCTTCTGT	1140		
Db	1081	TGAGTGGCTCTTCTGGGGCTGTGGGAGCTTTAGCTGTGGTGTGAGGTGCTTCTGT	1140		
Qy	1141	ATCCCCCAATATAGAGCCCACTTCAATCAATGTCTCAGGAGAAATGTGGACATGAGG	1200		
Db	1141	ATCCCCCAATATAGAGCCCACTTCAATCAATGTCTCAGGAGAAATGTGGACATGAGG	1200		
Qy	1201	ACTTTACTTGGGTACTCCACAGAGTACCTGTGGAGGGGTACAGAACTGGTTC	1260		
Db	1201	ACTTTACTTGGGTACTCCACAGAGTACCTGTGGAGGGGTACAGAACTGGTTC	1260		
Qy	1261	TGGGGGCCCCCGCTACAGCATACCGGAAAGCTGTCACTTTCAACAGGTGTCCAGGC	1320		
Db	1261	TGGGGGCCCCCGCTACAGCATACCGGAAAGCTGTCACTTTCAACAGGTGTCCAGGC	1320		
Qy	1321	AATGGAGAAAGAGCGAAGTCAAGGAGCGAGATGGCTCTTCTGGGGCTTCCC	1380		
Db	1321	AATGGAGAAAGAGCGAAGTCAAGGAGCGAGATGGCTCTTCTGGGGCTTCCC	1380		
Qy	1381	TCTCTCCGTGGATGTGGACAGATGGCAGACCGACCTGATCTCTATTTGGGGCCCCC	1440		
Db	1381	TCTCTCCGTGGATGTGGACAGATGGCAGACCGACCTGATCTCTATTTGGGGCCCCC	1440		
Qy	1441	ATTACTATGAGCAGACCGGAGGGGCGAGGTGTCCGTGTCCCTTGGGGGCGAGA	1500		
Db	1441	ATTACTATGAGCAGACCGGAGGGGCGAGGTGTCCGTGTCCCTTGGGGGCGAGA	1500		
Qy	1501	GGGTGCACTGGCAGTGTGACGCTTCTCCGTGGTGGAGAGGGCCACCTTGGGGCGCT	1560		
Db	1501	GGGTGCACTGGCAGTGTGACGCTTCTCCGTGGTGGAGAGGGCCACCTTGGGGCGCT	1560		
Qy	1561	TTGGGGCAGCCCTCACAGTGTGGGGATGTGAATGAGGACAAAGCTGATAGAGTGGCCA	1620		
Db	1561	TTGGGGCAGCCCTCACAGTGTGGGGATGTGAATGAGGACAAAGCTGATAGAGTGGCCA	1620		
Qy	1621	TTGGGGCCCCGGGAGAGCAGGAGAACCGGGGTCTGTCTACCTGTTTTCAGGAGCTCAG	1680		
Db	1621	TTGGGGCCCCGGGAGAGCAGGAGAACCGGGGTCTGTCTACCTGTTTTCAGGAGCTCAG	1680		
Qy	1681	AATCCGGCATCAGCCCTCCACAGCCAGCGGATTTGCCAGCTTCCCCAGGC	1740		
Db	1681	AATCCGGCATCAGCCCTCCACAGCCAGCGGATTTGCCAGCTTCCCCAGGC	1740		
Qy	1741	TGCAGTATTTTGGGAGCGCTGAGTGGGGGTGAGGACCTTCCAGAGTGGACTGATGG	1800		
Db	1741	TGCAGTATTTTGGGAGCGCTGAGTGGGGGTGAGGACCTTCCAGAGTGGACTGATGG	1800		
Qy	1801	ACCTGGCGTGGGGGCGGGGCGAGGTGTCTCTGCTCAGGAGTCTGCGGTGTGAAAG	1860		
Db	1801	ACCTGGCGTGGGGGCGGGGCGAGGTGTCTCTGCTCAGGAGTCTGCGGTGTGAAAG	1860		
Qy	1861	TGGGGGTGGCCATGAGATTTACGCCCTGTGGAGGTGGCCAGGGTGTGTACCGGTGTGGG	1920		
Db	1861	TGGGGGTGGCCATGAGATTTACGCCCTGTGGAGGTGGCCAGGGTGTGTACCGGTGTGGG	1920		
Qy	1921	AAGAGAGCCAGTGGCCCTGGAAGCTGGGAGCGCACCGTCTGTCTCACCATCCAGAAAA	1980		
Db	1921	AAGAGAGCCAGTGGCCCTGGAAGCTGGGAGCGCACCGTCTGTCTCACCATCCAGAAAA	1980		
Qy	1981	GCTCACTGGACAGCTAGGTGACATCCAAAGCTCTGTCAAGTTTGTGTGSCATGAGCC	2040		
Db	1981	GCTCACTGGACAGCTAGGTGACATCCAAAGCTCTGTCAAGTTTGTGTGSCATGAGCC	2040		
Qy	2041	CAGGTCTGTGACTTCTCGTCCCAATTTTCATGAACCAAGAACCCACCTTTGACTCGAA	2100		

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2004, 05:52:06 ; Search time 13747 Seconds
(without alignment)
8093.874 Million cell updates/sec

Title: US-09-891-943-1
Perfect score: 3726
Sequence: 1 tgaaccttggcactgtgttt.....agcataaaatttcataatgct 3736

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- EST:**
- 1: em_estba:**
 - 2: em_esthum:**
 - 3: em_estmu:**
 - 4: em_estmu:**
 - 5: em_estov:**
 - 6: em_estpl:**
 - 7: em_estro:**
 - 8: em_htc:**
 - 9: gb_est1:**
 - 10: gb_est2:**
 - 11: gb_htc:**
 - 12: gb_est3:**
 - 13: gb_est4:**
 - 14: gb_est5:**
 - 15: em_estfum:**
 - 16: em_estom:**
 - 17: em_gss_hum:**
 - 18: em_gss_inv:**
 - 19: em_gss_pln:**
 - 20: em_gss_vrt:**
 - 21: em_gss_fun:**
 - 22: em_gss_man:**
 - 23: em_gss_mus:**
 - 24: em_gss_pro:**
 - 25: em_gss_rod:**
 - 26: em_gss_phg:**
 - 27: em_gss_vzi:**
 - 28: gb_gss1:**
 - 29: gb_gss2:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1638.8	44.0	4056	11 BC042840	BC042840 Mus muscu
2	1404.4	37.7	3684	11 AK089521	AK089521 Mus muscu
3	912	24.5	3019	11 AK040921	AK040921 Mus muscu
4	844.8	22.7	3326	11 AK039444	AK039444 Mus muscu

5	488.4	13.1	1067	13	BQ711865
6	428	11.5	438	10	BF871208
7	391.6	10.5	918	13	BQ942456
8	384.6	10.3	572	12	BM366605
9	366.4	9.8	529	12	BM364453
10	356.4	9.6	659	13	BY739005
11	351.8	9.4	648	10	BB636941
12	351.2	9.4	545	12	BM366368
13	336.8	9.0	713	13	BY764281
14	326.2	8.8	1210	11	AK089167
15	319	8.6	658	10	BF135162
16	318.2	8.5	939	12	BF191010
17	311.4	8.4	753	10	BF160615
18	302.6	8.1	493	10	BF655065
19	301.2	8.1	908	13	BQ714983
20	300.4	8.1	819	10	BE571108
21	299	8.0	682	14	CS522420
22	295.4	7.9	491	10	BF601810
23	293.2	7.9	389	12	BI685527
24	292.8	7.9	389	12	BI679978
25	279.2	7.5	984	13	BQ708826
26	277	7.4	540	9	AI604269
27	273	7.3	546	9	AA178276
28	267	7.2	538	9	AI286783
29	264.2	7.1	666	13	BY749040
30	263	7.1	663	13	BY748444
31	263	7.1	689	13	BY749906
32	261.8	7.0	493	10	BF470350
33	260.4	7.0	684	13	BY749554
34	260.2	7.0	386	10	AW141017
35	259.8	7.0	692	13	BY748124
36	259.4	7.0	700	13	BY750010
37	258.4	6.9	482	12	BI683165
38	255.8	6.9	702	13	BY748616
39	254.6	6.8	678	13	BY749866
40	254.2	6.8	372	10	AW446046
41	253.8	6.8	805	12	BI645825
42	252.2	6.8	376	12	BM146659
43	252.2	6.8	722	13	EX644572
44	250.8	6.7	656	13	BY747704
45	248.4	6.7	803	12	BI107542

ALIGNMENTS

RESULT 1
BC042840
LOCUS Mus musculus, Similar to integrin alpha M; clone IMAGE:3988275, 4056 bp mRNA linear HTC 16-JAN-2003
DEFINITION BC042840
ACCESSION BC042840
VERSION BC042840.1 GI:27768957
KEYWORDS HTS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 4056)
Strausberg, R.
AUTHORS Direct Submission
TITLE Submitted (02-JAN-2003) National Institutes of Health, Mammalian
JOURNAL Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
REMARK Contact: MGC help desk
COMMENT Email: cgabbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center

Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: amg@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulsegod, H.,
 Kowis, C.R., Sneed, A.O., Martin, R.G., Muzny, D.M., Nanavati,
 A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAP Plate: 30 Row: 9 Column: 12
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 10946645
 This clone has the following problem: frame shifted.

FEATURES

source

1..4056
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C2ECH II"
 /db_xref="taxon:10090"
 /clone="IMAGE:3988275"
 /tissue_type="Mammary tumor metastasized to lung. Tumor
 arose spontaneously from a senescent normal mammary
 (clonal) outgrowth infected with the virus MMTV."
 /clone_lib="NCI CGAP_Lu29"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"

ORIGIN

Query Match 44.0%; Score 1638.8; DB 11; Length 4056;
 Best Local Similarity 68.0%; Pred. No. 0;
 Matches 2363; Conservative 0; Mismatches 1072; Indels 41; Gaps 4;

QY 3 ACCTTCGGACCTGCTCTCTGAGTGCCTGCTCTTATCATGAGATTCAACTGGAT 62
 DB 63 ACCTGGATAGCTTTCTTCTGCTGTGGGTTGTTTCTTGTCTTGGCTTCACTGGAT 122
 QY 63 GTGAGGAGCTACGATCTTCAGGAGGATGAGCGGCTTTGGGCAAGCGTGTGAG 122
 DB 123 GCAGAGAGCTGACATTTTACATGAGCGGTGCTGAGTTCGGACACAGTGTGCTCCAG 182
 QY 123 TTGGTGGATCTCGACTCGTGTGGGAGCACCCCTGGAGGTGCTGCGGCGCAACAGAGC 182
 DB 183 TATGATAGTTCCTGGTGTGGTGTGGAGCACCAGAGAAATAAAGCCACTACCAATA 242
 QY 183 CGAGGCTGTATGACTGCGCAGCTGCCACCGCATGTGCCAGCCATCCCGTGCACATC 242
 DB 243 GTTGGCTCTACAAATGTGCTATACACAGGCACTGTGAGCCCATCTCCCTCCAGTG 302
 QY 243 CGCCCTGAGCGCTGAACATGCTTGGGCTGACCCCTGGCAGCTCCACCAACGGCTCC 302
 DB 303 CCCCAGAGGCTGTGAACATGCTCCCTGGGCTGTCCCTTGTGCTGCGCACCAACCTTCC 362
 QY 303 CGGCTCTGGCTGTGGCCCGGAGCTGTGACAGAGTCTGTGGGAGAACTCATACTCAAG 362
 DB 363 TGGCTGTGGCTGTGGCTTCTACTGTGACACACATGACAGAGAAATATATCTGACA 422
 QY 363 GGTTCCTGCTCTGCTGGCTCGGCTGGGAGATCATCCAGACAGTCCCGAGCCAGC 422
 DB 423 GG--GCTGCTTTCTACTGATTCATCATTCAGCAGAGCCAGAACTTCCCAACTGACAG 480
 QY 423 CCAGAGTGTCCATCAAGAGATGGACATCGTCTTCTGATTGACGGCTCTGGAAGCAATT 482
 DB 481 CAGAGTGTCCAAAGAGACAGACATGCTGTTCTGATTGATGGCTCGGCTAGCATC 540
 QY 483 GACCAAAATGACTTTTAAACAGATGAAGGCTTTGTGTCAGCTGTCTATGGCCAGTTTGA 542
 DB 541 AGTTCCACAGATTTTGAATAAATGCTGCACTTTGTTAAAGCTGTGATGAGCCAGCTTCAG 600
 QY 543 GGCAGTGCACCTGTTTGGCTGATGAGTCACTCAAACTCTCTGAAGATCCACTTCAAC 602
 DB 601 AGACCTAGACACAGGTTCTCCCTGATGAGCTTCTCTGATTTACTTCCAGTACATTTTACT 660
 QY 603 TTCACCAATTCGGAGCAGCCCGAGCCAGAGAGCGCTGTGGATCCCATCGTCCAACTG 662

DB 661 TTCAACAACCTTCTCATCTCCAGCTCAAGGCTTTAAGTCTGTTGGATCTCTGTAAGCAGCTA 720
 QY 663 AAAGGCTGACGTTTCAACGCGCCACCGGCAATCTCTGACAGTGGTGACACAGCTATTTTCATCAT 722
 DB 721 AGAGGGTACACATACACAGCTCTGGCTATCAAGCATGTCTAACAAGAACTGTCTTCCACCAC 780
 QY 723 AGAATGGGCGCCGAAAGTGCCTCAAGAGATCTCTATTGTCTATCAGATGAGGAGAG 782
 DB 781 CAAAGTGGAGCTCTGGCAGATGCCAAGGTCTCTCATTTGTCATCTCTGATGGGAGAAA 840
 QY 783 TACAAAGACCCCTCGAATACAGTATGTATCTCCCGAGGAGAGAAAGCTGGCATCTATC 842
 DB 841 CAAGGGGCAACCTTGAGTTATGATGTATCTATCTCCCATGCGAGAGGCTGCAAGCATCAT 900
 QY 843 CGCTAGCTATCGGGTGGGACACGCTTTCAGGAGCCACCTGCGCAGGAGGCTGAAT 902
 DB 901 CGTTATGCAATTTGGGTAGGAAGGCCCTTTTAAATGAACATCTCCAGCAAGATTAAG 960
 QY 903 ACCATCAGCTCAGCGCTCCCGAGGACCACTGTTTCAAGGTGGACAACTTTTCAGCCCTT 962
 DB 961 GCATTTGCATCGATGCTTCCCATGAATACGTATTACGGGTGGAGAACTTTTGTATGCTTTG 1020
 QY 963 GGAGAGTCCAGAGCAGCTGAGGAGAGATCTATGCAAGTTGAGGAAACCCAGTCCAGG 1022
 DB 1021 AAGGATATTGAGAAATCAGCTGAAGGAGAAAGTCTTTGCCATTTGAGGGCAGAGACCA 1080
 QY 1023 GCAAGCAGCTCTTCCAGCAGCAGATGTCCCAAGAGGCTTCAGCAGACCCCTCACAATG 1082
 DB 1081 AGCAGCAGTACTTTTGAATTTGGAGATGTCCAGGAGGCTTCAGTCTGTGTTTACACCT 1140
 QY 1083 GATGGGCTCTTCTCGGGGCTGTGGGAGCTTTAGTGTCTGGAGGTGCTTCTGTTAT 1142
 DB 1141 GATGAGCAGCTTCTGGGGGCTGTGGGAACTTCAGCTGTCTGGAGGTGCTTCTTGTATC 1200
 QY 1143 CCCCCAAATATAGCCCCCACTTCAACAATGTCTCAGGAGAAATGTGACATGAGGAGC 1202
 DB 1201 CCTCAAAATATAGACCCACCTTCAACAATGTCTCAGGAGAACCCAGCAGATGAGGAGC 1260
 QY 1203 TCTTACCTGGGTTACTCCACCGAGCTAGGCTGTGGAAGGGGTGACAGAACCTGTGCTG 1262
 DB 1261 GCTTACCTGGGTTACTCCACCGCAGCTGGGCTTTTGAAGGGGTCCACAGCTGTGCTG 1320
 QY 1263 GGGGCCCCCGCTTACAGCATACCGGAGGCTGTCTATCTTACCAGGTGTCCAGGCA 1322
 DB 1321 GGGGCCCCCTCGCCACACAGCAGCGGAGGTTGTCTATCTTACCAGGAATCCAGGAC 1380
 QY 1323 TGGAGGAAAGGCCGGAAGTCAACAGGACCGCATCGGCTCTCTTCTGGGGCTCTCCCTC 1382
 DB 1381 TGGAGGCCCAAGTCTGAAGTCAGAGGACACAGATCGGCTCTCTTCTGGGGCTCTCTT 1440
 QY 1383 TGCTCCGTGATGTGACAGGATGGCAGCAGCATGATCTCTCATTTGGGGGCCCCCAT 1442
 DB 1441 TGTCTGTGACATGTGATAGATGGCAGCAGCTGCTGCTGTGATTTGAGTCCCCCAT 1500
 QY 1443 TACTATGAGCAGACCCCGAGGGGSCAGGTGTCTGCTGCTCTCTTGGCTAGGGGCGAGG 1502
 DB 1501 TACTATGAGCAGACCCCGAGGGGSCAGGTGTCTGAGTGTGCCCCATGCCCGGTGGGAG- 1559
 QY 1503 GTGAGTGGCAGTGTGAGCTGTCTCCGTGGTGGAGGAGGAGCCACCCCTGGGGCGGCTTT 1562
 DB 1560 --CAGGTGGCAATTGTGGGACCACTCCATGGGAGCAGGGGCCATCTTGGGGCGGCTTT 1617
 QY 1563 GGGGCGAGCCCTGACAGTGTGTTGGGGGATGTGAATGAGGACAAAGCTGTAGAGCTGGCCATT 1622
 DB 1618 GGGGCGGCTCTGACAGTGTAGGGGACGTGAATGGGAGCAGTCTGGCAGATGTGGCTATT 1677
 QY 1623 GGGGCCCCCGGAGAGCAGAGAGAACCGGGGTGCTGTCTACCTGTTTTCAGGAGCCCTCAGAA 1682
 DB 1678 GGTGACCCCGGAGAGGAGGAGAACAGAGGTGCTGTCTACATATTTTCATGAGCCCTCAGA 1737
 QY 1683 TCCGSCATCAGCCCTCCACAGCAGCGGATTCGACAGCTCCCAAGCTCTCCCCAGGCTG 1742

xx A probe based on a partial cDNA clone (given in AQ09127) of canine alpha
 CC -TM1 was used to screen a human spleen cDNA library to identify clone
 CC 19A2 encoding the beta-1 integrin alpha-subunit, alpha-d. The cDNA was
 CC manipulated to allow expression of recombinant alpha-d subunit in COS and
 CC CHO cells

XX	Sequence	3726 BP	817 A	1054 C	1054 G	791 T	0 U	0 Other
Qy	Query Match	100.0%	Score 3726	DB 2	Length 3726			
Qy	Best Local Similarity	100.0%	Pred. No. 0					
Qy	Matches 3726	Conservative 0	Mismatches 0	Indels 0	Gaps 0			
Qy	1	TGACCTTCGGCACTGCTCTTCTTCAGTGTCTCGCTCTTATCATGATTCACACTGG	60					
Db	1	TGACCTTCGGCACTGCTCTTCTTCAGTGTCTCGCTCTTATCATGATTCACACTGG	60					
Qy	61	ATGTGAGGAGCTACGATCTCCACGAGGATSCAGCGCTTTGGGCGAGAGCTGTGC	120					
Db	61	ATGTGAGGAGCTACGATCTCCACGAGGATSCAGCGCTTTGGGCGAGAGCTGTGC	120					
Qy	121	AGTTTCGGTGTACTTCGACTCGTGTGGGACACCCCTCGAGGTGTGGCGGCCAACACGA	180					
Db	121	AGTTTCGGTGTACTTCGACTCGTGTGGGACACCCCTCGAGGTGTGGCGGCCAACACGA	180					
Qy	181	CGGACCGCTGTATGACTCGGACGCTGCCACCGGATGTGCCAGCCATCCCGCTGCACA	240					
Db	181	CGGACCGCTGTATGACTCGGACGCTGCCACCGGATGTGCCAGCCATCCCGCTGCACA	240					
Qy	241	TCGGCCCTGAGGCGGTGAACATGTCTTTGGGCTGACCTTCGAGCGCTCCACCAACGGCT	300					
Db	241	TCGGCCCTGAGGCGGTGAACATGTCTTTGGGCTGACCTTCGAGCGCTCCACCAACGGCT	300					
Qy	301	CCGGCTTCCTGCGCTGTGGCCGACCTGCACAGAGTCTGTGGGAGAACTCATCTCAA	360					
Db	301	CCGGCTTCCTGCGCTGTGGCCGACCTGCACAGAGTCTGTGGGAGAACTCATCTCAA	360					
Qy	361	AGGTTCTCTGCTCTGCTGGGCTCGGCTGGGAGATCATCCAGACAGTCCCCGACGCCA	420					
Db	361	AGGTTCTCTGCTCTGCTGGGCTCGGCTGGGAGATCATCCAGACAGTCCCCGACGCCA	420					
Qy	421	CGCCAGAGTGTCCACATCAAGAGATGACATCGTCTTCCTGTGATGACGGCTTGGGAACA	480					
Db	421	CGCCAGAGTGTCCACATCAAGAGATGACATCGTCTTCCTGTGATGACGGCTTGGGAACA	480					
Qy	481	TTGACCAAAATGACTTTAAACAGATGAAGGCTTTGTCACAGCTGTCTATGGGCGAGTTG	540					
Db	481	TTGACCAAAATGACTTTAAACAGATGAAGGCTTTGTCACAGCTGTCTATGGGCGAGTTG	540					
Qy	541	AGGCACTGACACCTCTTTGCACTGTAGTAGTACTCAAAACCTCTGAGATCCACTTCA	600					
Db	541	AGGCACTGACACCTCTTTGCACTGTAGTAGTACTCAAAACCTCTGAGATCCACTTCA	600					
Qy	601	CTTTACCCCAATTCGGACACCGCCGACGACGAGCGCTGTGGATCCCATCGTCCAC	660					
Db	601	CTTTACCCCAATTCGGACACCGCCGACGACGAGCGCTGTGGATCCCATCGTCCAC	660					
Qy	661	TGAAGGCTCAGCTTTCACGGCCACGGGCTCTCTGACAGTGTGGACACAGCTATTTTCATC	720					
Db	661	TGAAGGCTCAGCTTTCACGGCCACGGGCTCTCTGACAGTGTGGACACAGCTATTTTCATC	720					
Qy	721	ATAAGATGGGGCCGAAAAGTGCAGAGATCCTCATGTGTATCATCAGATGGGCGAGA	780					
Db	721	ATAAGATGGGGCCGAAAAGTGCAGAGATCCTCATGTGTATCATCAGATGGGCGAGA	780					
Qy	781	AGTCAAGACCCCTCGGAATACAGTGTATGTATCCCGCAGGACAGAAAGGCTGGCATCA	840					
Db	781	AGTCAAGACCCCTCGGAATACAGTGTATGTATCCCGCAGGACAGAAAGGCTGGCATCA	840					
Qy	841	TCGGTACCGTATCGGGGTGGGACACGGTTTCCAGGACCCCATCTGCCAGCAGAGCTGA	900					
Db	841	TCGGTACCGTATCGGGGTGGGACACGGTTTCCAGGACCCCATCTGCCAGCAGAGCTGA	900					
Qy	901	ATACCATCAGCTCAGCGCTCCGAGGACACGCTGTTCAGGTTGGACAACACTTTCAGCCCC	960					
Db	901	ATACCATCAGCTCAGCGCTCCGAGGACACGCTGTTCAGGTTGGACAACACTTTCAGCCCC	960					
Qy	961	TTGGACGATCCAGAGCAGCTGCAGGAGAAAGATCTATGCAAGTTGAGGGAAACCCAGTCCA	1020					

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2004, 05:52:06 ; Search time 282 Seconds
(without alignments)
7332.436 Million cell updates/sec

Title: US-09-891-943-1
Perfect score: 3726
Sequence: 1 tgccttcggcactgtgctt.....agcataatttcattatgct 3726

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/ECTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3726	100.0	3726	1 US-08-173-497-1	Sequence 1, Appli
2	3726	100.0	3726	1 US-08-286-889-1	Sequence 1, Appli
3	3726	100.0	3726	1 US-08-485-618-1	Sequence 1, Appli
4	3726	100.0	3726	1 US-08-362-652-1	Sequence 1, Appli
5	3726	100.0	3726	1 US-08-605-672-1	Sequence 1, Appli
6	3726	100.0	3726	2 US-08-482-293A-1	Sequence 1, Appli
7	3726	100.0	3726	2 US-08-943-363-1	Sequence 1, Appli
8	3726	100.0	3726	3 US-09-193-043-1	Sequence 1, Appli
9	3726	100.0	3726	4 US-09-688-307A-1	Sequence 1, Appli
10	3726	100.0	3726	4 US-09-350-259-1	Sequence 1, Appli
11	3699	99.3	3785	1 US-08-485-618-98	Sequence 98, Appli
12	3699	99.3	3785	1 US-08-605-672-98	Sequence 98, Appli
13	3699	99.3	3785	2 US-08-482-293A-98	Sequence 98, Appli
14	3699	99.3	3785	2 US-08-943-363-98	Sequence 98, Appli
15	3699	99.3	3785	3 US-09-193-043-98	Sequence 98, Appli
16	3699	99.3	3785	4 US-09-688-307A-98	Sequence 98, Appli
17	3699	99.3	3785	4 US-09-350-259-98	Sequence 98, Appli
18	3583	96.2	3956	1 US-08-485-618-97	Sequence 97, Appli
19	3583	96.2	3956	1 US-08-605-672-97	Sequence 97, Appli
20	3583	96.2	3956	2 US-08-482-293A-97	Sequence 97, Appli
21	3583	96.2	3956	2 US-08-943-363-97	Sequence 97, Appli
22	3583	96.2	3956	3 US-09-193-043-97	Sequence 97, Appli
23	3583	96.2	3956	4 US-09-688-307A-97	Sequence 97, Appli
24	3583	96.2	3956	4 US-09-350-259-97	Sequence 97, Appli
25	2315.2	62.1	3803	1 US-08-485-618-52	Sequence 52, Appli
26	2315.2	62.1	3803	1 US-08-362-652-52	Sequence 52, Appli
27	2315.2	62.1	3803	1 US-08-605-672-52	Sequence 52, Appli

28	2315.2	62.1	3803	2 US-08-482-293A-52	Sequence 52, Appli
29	2315.2	62.1	3803	2 US-08-943-363-52	Sequence 52, Appli
30	2315.2	62.1	3803	3 US-09-193-043-52	Sequence 52, Appli
31	2315.2	62.1	3803	4 US-09-688-307A-52	Sequence 52, Appli
32	2315.2	62.1	3803	4 US-09-350-259-52	Sequence 52, Appli
33	2304.2	61.8	3597	1 US-08-485-618-54	Sequence 54, Appli
34	2304.2	61.8	3597	1 US-08-362-652-54	Sequence 54, Appli
35	2304.2	61.8	3597	1 US-08-605-672-54	Sequence 54, Appli
36	2304.2	61.8	3597	2 US-08-482-293A-54	Sequence 54, Appli
37	2304.2	61.8	3597	2 US-08-943-363-54	Sequence 54, Appli
38	2304.2	61.8	3597	3 US-09-193-043-54	Sequence 54, Appli
39	2304.2	61.8	3597	4 US-09-688-307A-54	Sequence 54, Appli
40	2304.2	61.8	3597	4 US-09-350-259-54	Sequence 54, Appli
41	2300.2	61.7	3528	1 US-08-286-889-36	Sequence 36, Appli
42	2300.2	61.7	3528	1 US-08-485-618-36	Sequence 36, Appli
43	2300.2	61.7	3528	1 US-08-362-652-36	Sequence 36, Appli
44	2300.2	61.7	3528	1 US-08-605-672-36	Sequence 36, Appli
45	2300.2	61.7	3528	2 US-08-482-293A-36	Sequence 36, Appli

ALIGNMENTS

RESULT 1
US-08-173-497-1
; Sequence 1, Application US/08173497
; Patent No. 5437958
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van Der Vieren, Monica
; TITLE OF INVENTION: No. 5437958el Human 2 Integrin Alpha
; TITLE OF INVENTION: Subunit
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 S. Wacker Drive, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/173,497
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5437958and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/31363
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3726 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3..3485
US-08-173-497-1

Query Match 100.0%; Score 3726; DB 1; Length 3726;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3726; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2004, 09:36:53 ; Search time 27 Seconds
(without alignments)
4136.236 Million cell updates/sec

Title: US-09-891-943-2

Perfect score: 5987

Sequence: 1 TFGTVLLSLVLSYHGFGFLD.....DTATFGDDPFCVAPNVPLS 1161

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Pir1.*

2: Pir2.*

3: Pir3.*

4: Pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3818	63.8	1163	1 RWHU1C	cell surface glyco
2	3455	57.7	1153	1 RWHU1B	cell surface glyco
3	3327.5	55.6	1153	2 S00551	leukocyte surface
4	1474	24.6	1170	2 S03308	cell surface glyco
5	1457.5	24.3	1163	2 I56126	lymphocyte fuction
6	1148	19.2	1179	2 A53213	integrin alpha-E c
7	1030.5	17.2	1151	2 A52226	integrin alpha-1 c
8	1030.5	17.2	1180	2 A35854	integrin alpha-1 c
9	1012	16.9	1178	2 S44142	VLA-2 protein homo
10	1007.5	16.8	1170	2 I45914	integrin alpha 2 B
11	1007.5	16.8	1181	2 A33998	integrin alpha-2 c
12	654.5	10.9	1038	2 S06046	integrin alpha-4 c
13	646.5	10.8	1039	2 A41131	lymphocyte-Peyer's
14	638	10.7	1054	2 JG7294	alhap integrin -
15	611	10.2	1041	2 T31437	integrin alpha cha
16	584.5	9.8	1035	2 I58409	integrin alpha-9 c
17	579.5	9.7	1053	2 I55534	VLA-3 alpha subuni
18	576.5	9.6	1051	2 A35761	cell surface glyco
19	555	9.3	1037	2 A60163	glycoprotein IIB -
20	552	9.2	1049	2 A27079	fibronectin recept
21	548	9.2	1039	2 A34269	integrin alpha-2b
22	547.5	9.1	1051	2 A40021	integrin alpha-3 c
23	535.5	8.9	1053	2 S44250	integrin alpha-5 c
24	534	8.9	1034	2 A36108	integrin alpha-V c
25	524.5	8.8	1394	2 A29637	position-specific
26	497	8.3	1044	2 T10050	integrin alpha-v c
27	496.5	8.3	1072	2 A38457	integrin alpha-6 c
28	493	8.2	1048	2 A27421	integrin alpha-5 c
29	492	8.2	1146	2 S40311	integrin - fruit f

30 477.5 8.0 1044 2 S16516 integrin alpha-8 c
31 473 7.9 1073 2 B36429 integrin alpha-6 c
32 464 7.8 1137 2 JCS950 integrin alpha-7 c
33 462.5 7.7 1091 2 A41543 integrin alpha-6 c
34 446.5 7.5 1045 2 S60571 integrin alpha v c
35 439 7.3 1135 2 I81186 alpha-7 integrin -
36 433.5 7.2 1115 2 T09433 integrin alpha cha
37 431.5 7.2 1115 2 T09403 integrin alpha cha
38 424.5 7.2 1226 2 S44824 F54F2.1 protein -
39 420.5 7.0 1106 2 S38783 integrin alpha cha
40 410.5 6.9 764 2 I36916 glycoprotein IIB -
41 394 6.6 1139 2 S28277 hypothetical prote
42 322.5 5.4 272 2 A55348 integrin alpha-1 -
43 319.5 5.3 126 2 B30892 leukocyte adhesio
44 319.5 5.3 1086 2 T18523 integrin alpha cha
45 313.5 5.2 604 2 I36917 glycoprotein IIB -

ALIGNMENTS

RESULT 1

RWHU1C

cell surface glycoprotein CD11c precursor - human

N;Alternate names: leukocyte adhesion receptor p150,95 alpha chain

C;Species: Homo sapiens (man)

C;Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text_change 22-Jun-1999

C;Accession: A36584; A35543; S00864

R;Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.

J. Biol. Chem. 265, 12750-12751, 1990

A;Reference number: A36584

A;Contents: erratum

A;Accession: A36584

A;Molecule type: DNA

A;Residues: 1-1163 <COR>

A;Note: this revision to the sequence from reference A35543 includes the carboxyl end
R;Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.

J. Biol. Chem. 265, 2782-2788, 1990

A;Title: Genomic structure of an integrin alpha subunit, the leukocyte p150,95 molecule.

A;Reference number: A35543; MUID:90153906; PMID:2303426

A;Accession: A35543

A;Molecule type: DNA

A;Residues: 1-834 <CO2>

A;Note: this sequence has been revised in reference A36584

R;Corbi, A.L.; Miller, L.J.; O'Connor, K.; Larson, R.S.; Springer, T.A.

EMBO J. 6, 4023-4028, 1987

A;Title: cDNA cloning and complete primary structure of the alpha subunit of a leukocyte

A;Reference number: S00864; MUID:88166645; PMID:3327687

A;Accession: S00864

A;Molecule type: mRNA

A;Residues: 1-755,'L',757-1163 <CO3>

A;Cross-references: GB:M81695; EMBL:Y000093; NID:9487829; PIDN:AAA59180.1; PID:9487830

A;Note: part of this sequence was confirmed by protein sequencing

A;Comment: A common beta chain (CD18) forms a heterodimer with CD11b to form Mac-1 on mye

C;Genetics:

A;Gene: GDB:ITGAX; CD11C

A;Cross-references: GDB:119758; OMIM:151510

A;Map position: 16p11.2-16p11.2

C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homol

C;Keywords: calcium; cell adhesion; glycoprotein; heterodimer; magnesium; tandem repeat;

F;1-19/Domain: signal sequence #status predicted <SIG>

F;20-1163/Product: cell surface glycoprotein CD11c #status predicted <MAT>

F;20-1107/Domain: extracellular #status predicted <EXT>

F;149-319/Domain: von Willebrand factor type A repeat homology <VWA4>

F;1108-1133/Domain: transmembrane #status predicted <TM>

F;1134-1163/Domain: intracellular #status predicted <INT>

F;61,89,392,697,735,899,939,1050/Binding site: carbohydrate (Asn) (covalent) #status prec

Query Match 63.8%; Score 3818; DB 1; Length 1163;

Best Local Similarity 65.5%; Pred. No. 6,6e-267;

Matches 755; Conservative 138; Mismatches 253; Indels 6; Gaps 5;

QY 1 TFGTVLLSLVLSYHGFGFLDVEEPTIFQDAGFGQSVVQFGSRLVWGAPLEWAAQT 60

Db 4 TRALLLPTALATSGFNLDTBELTAFVDSAGFGDSVVOYANSVVVVGAPQKITAANQT 63
QY 61 GRLYDCAATGWCQPIPLHPIPEAVNMSLGLTAASTNGSRLACGPTLHRVCGENSYSK 120
Db 64 GGLYQCGYSTGANCEPIGLQVPEAVNMSLGLSLASTTSPSOLLACGPTVHHECGRNWYLT 123
QY 121 GSCLLGSRWELIQVPDAPTECPCHQEMDMIVFLIDGSGSIDQNDPKNQKGVQAVMGQFE 180
Db 124 GLCFLLGPT-QLTORLPVSROCEPQEDIVFLIDGSGSISRNPATFATWNNFVRAVISQFQ 182
QY 181 GDTTLFALMOYNSNLKIHFTFTQFTSFSQOQLVDPIVOLKGLTFTAGILTVTQLFHH 240
Db 183 RPSTQFSLMOFSNKQFTHTFEFRFTSNPLSLASVHQLQGPVITATAIQNVVHRLFHA 242
QY 241 KNGARKSAKILIVITDGOKYKDPLEYSDVPOAEKAGIIRYAGVGHAFQGPARTARQELN 300
Db 243 SYGARRDATKILIVITDGKKGSDSDYKDVIPMDAAGIIRYAGVGLAFQNRNSWKELN 302
QY 301 TISSAPPOHVPKVNFAALGSIQKOLGKIYAVGTSOGRASSSSFOHEMSQEGFSTALTM 360
Db 303 DIASKPSQSHIFKVEDFDLKDQIQNLKEKIFPAIGETTTSSSSFELEWAQEGFSAVPTP 362
QY 361 DGLFLGAVGSFWSGAFILYPPNMGPTFINMSQENVMDRDSVYLGYSYTELALWKGVQNLVL 420
Db 363 DGPVLGAVGSFTWSGAFILYPPNMGPTFINMSQENVMDRDSVYLGYSYTELALWKGVQSLVL 422
QY 421 GAPRYOHTKAVITQVSRQWKKAEVTQIGSYFGASLCSVDVDSGSDTLILIGAPH 480
Db 423 GAPRYOHTKAVITQVSRQWKKAEVTQIGSYFGASLCSVDVDSGSDTLILIGAPH 482
QY 481 YTEQTRGGQVSVCLPRGQORVQWQCDVLRGEGHPWCRFGAALTVLGVDNEDKLIIDVAI 540
Db 483 YTEQTRGGQVSVCLPRGWR-RWMCDAVLYGEGHPWCRFGAALTVLGVDNEDKLIIDVAI 541
QY 541 GAPGEENRGAVYLFHGAESGISPSHSQRIASSQLSPLOYFGALSGQDLTQDGLMD 600
Db 542 GAPGEENRGAVYLFHGVLPSPHSQRIAGSLSLSELPFGQALSGQDLTQDGLMD 601
QY 601 LAVGARGQVLLRLSLPVLKGVAMRSPVEKAVYRCWEEKPSALEAGDAVTCITIOKS 660
Db 602 LAVGARGQVLLRLTRPVLWVGSMQIPAEIPRSAFECEQVQVSEQLTVQSNICLVIDKR 661
QY 661 SLDOLG--DIQSVRRFDLALDQRLTSRAIPNETKPTTRKTGLGTHCEFTKILLAPD 718
Db 662 SKNLLGSRDIQSSVTLTDLALDQRLSPRATFQETKRSLSRVRLVGLKACHENFNLLPFS 721
QY 719 CVDVSVPIILHNLNPSLREPIPSQNLRLPVLAVGSDQLFTASLPPEKNCQDGLCEGDL 778
Db 722 CVDVSVPIILNPLVGLKPLAFNLRLPMLAADAQRYFTASLPPEKNCQDGLHICQDNL 781
QY 779 GVTLSPSGQLTLVGSLSLENLVITVWVNGEDSYGTWLSLYPAGLSHRVRSQAQKPHQ 838
Db 782 GISFSPGLKSLVGNLNAEVMVWVNDGEDSYGTITFTSPAGLSYRYVAEGQKQOL 841
QY 839 SALRLACETVPTDEGLSRSCSNVHPIFEGSGNFTIVTFDVSYSKATLGDRMLKRAAS 898
Db 842 RSLHLTCDSPVSGQVTSWTSCHNLIFPGGAQITFLATFVSPKAVLGDRLLTANYS 901
QY 899 SENKASSKATQLELPVKYAVYTMISROESTKYPNATSEKMKKAEHRYRVNLS 958
Db 902 SENNTPRTSKTTFQLELPVKYAVYTVSSHEQTKYLNFSSEKESHVAMHRYVNNLG 961
QY 959 QRDLAISINFWPVLNGVAVM-DVWEAP-SQSLPCVSEKPPQSDFLTQISRSPLMD 1016
Db 962 QRDLPVFINFWPVELNQEAVMVMDVESHVQNFQNSLRCSSEKIPAPSDFLAHKQKPVLD 1021
QY 1017 CSIADCLQFRCDVPSFVSQBELDFTLXGNLSFGVRETLOKKVLVSVABITFTSVYSQ 1076
Db 1022 CSIAGCLRFRCVPSFVSQBELDFTLXGNLSFGVWRQILQKVSWSVABITFTSVYSQ 1081
QY 1077 LPQGEAFWRAQWVLEDEVDVNAIPIMSSVCALLILALITATYKLGFFKRVHVKEML 1136

Db 1082 LPQGEAFWRAQWVLEDEVDVNAIPIMSSVCALLILALITATYKLGFFKRVHVKEML 1141
QY 1137 EDKPEDTATFSG 1148
Db 1142 BEANGQIAPENG 1153
RESULT 2
RWVUB
cell surface glycoprotein CD11b precursor [validated] - human
N;Alternate names: complement receptor type 3 alpha chain; leukocyte adhesion protein Mac
eukocyte integrin alpha chain; neutrophil adherence receptor alpham chain
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text change 31-Dec-2000
C;Accession: A31108; A28915; A41600; A30892; A32218; A46526; A26091; I52567
R;Corbi, A.L.; Kishimoto, T.K.; Miller, L.J.; Springer, T.A.
J. Biol. Chem. 263, 12403-12411, 1988
A;Title: The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, CD11b).
A;Reference number: A31108; MUID:88315033; PMID:2457584
A;Accession: A31108
A;Molecule type: mRNA
A;Residues: 1-1153 <COR>
A;Cross-references: GB:J03925; NID:gl87284; PIDN:AAA59544.1; PID:G3071148
A;Note: part of this sequence was confirmed by protein sequencing
R;Arnaout, M.A.; Gupta, S.K.; Pierce, M.W.; Tenen, D.G.
J. Cell Biol. 106, 2153-2158, 1988
A;Title: Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor M.
A;Reference number: A28915; MUID:88257215; PMID:2454931
A;Accession: A28915
A;Molecule type: mRNA
A;Residues: 1-499,501-965, 'P', 967-1153 <ARN>
A;Cross-references: GB:M18044; GB:J03270; GB:M19664; GB:X07421; NID:gl86935; PIDN:AAA5945
A;Note: the authors translated the codon TAC for residue 1129 as Thr
A;Note: part of this sequence, including the amino end of the mature protein, was confirm
R;Shelley, C.S.; Arnaout, M.A.
Proc. Natl. Acad. Sci. U.S.A. 88, 10525-10529, 1991
A;Title: The promoter of the CD11b gene directs myeloid-specific and developmentally regu
A;Reference number: A41600; MUID:92073319; PMID:1683702
A;Accession: A41600
A;Molecule type: DNA
A;Residues: 1-9 <SHE>
A;Cross-references: GB:M76724; NID:gl80018; PIDN:AAA58410.1; PID:G553215
R;Arnaout, M.A.; Remold-O'Donnell, E.; Pierce, M.W.; Harris, P.; Tenen, D.G.
Proc. Natl. Acad. Sci. U.S.A. 85, 2776-2780, 1988
A;Title: Molecular cloning of the alpha-subunit of human and guinea pig leukocyte adhesic
A;Reference number: A94193; MUID:88190151; PMID:2833753
A;Accession: A30892
A;Molecule type: mRNA
A;Residues: 917-1042 <AR2>
A;Cross-references: GB:M18044
R;Hickstein, D.D.; Hickey, M.J.; Ozols, J.; Baker, D.M.; Back, A.L.; Roth, G.J.
Proc. Natl. Acad. Sci. U.S.A. 86, 257-261, 1989
A;Title: cDNA sequence for the alpham subunit of the human neutrophil adherence receptor
A;Reference number: A32218; MUID:89098893; PMID:2563162
A;Accession: A32218
A;Molecule type: mRNA
A;Residues: 9-1153 <HIC>
A;Cross-references: GB:J04145; NID:gl89068; PIDN:AAA59903.1; PID:G386975
A;Note: part of this sequence was confirmed by protein sequencing
R;Fleming, J.C.; Fahl, H.L.; Gonzalez, D.A.; Smith, T.F.; Tenen, D.G.
J. Immunol. 150, 480-490, 1993
A;Title: Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-int
n during evolution.
A;Reference number: A46526; MUID:93123748; PMID:8419480
A;Accession: A46526
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-499,501-1153 <PLE>
A;Cross-references: GB:S52227; NID:G263047; PIDN:AA24821.1; PID:G263049
A;Note: the last three bases of intron 13, CAG, are included in some but not all mature n
A;Note: sequence extracted from NCBI backbone (NCBIP:121963)
R;Pierce, M.W.; Remold-O'Donnell, E.; Todd III, R.F.; Arnaout, M.A.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2004, 09:33:03 ; Search time 17 Seconds
(without alignments)
3556.085 Million cell updates/sec

Title: US-09-891-943-2

Perfect score: 5987
Sequence: 1 TFGTVLLSLVSLVHGFNLD.....DTATFSGDDFSCVAPNPVLS 1161

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5987	100.0	1162	1 ITAD_HUMAN	Q13349 homo sapien
2	3815	63.7	1163	1 ITAX_HUMAN	P20702 homo sapien
3	3439.5	57.4	1152	1 ITAM_HUMAN	P11215 homo sapien
4	3327.5	55.6	1153	1 ITAM_MOUSE	P05555 mus musculus
5	1476	24.7	1170	1 ITAL_HUMAN	P20701 homo sapien
6	1457.5	24.3	1163	1 ITAL_MOUSE	P24063 mus musculus
7	1156.5	19.3	1167	1 ITAE_MOUSE	Q60677 mus musculus
8	1148	19.2	1179	1 ITAE_HUMAN	P38570 homo sapien
9	1146.5	19.1	1189	1 ITAH_HUMAN	Q9UKXS homo sapien
10	1131.5	18.9	1167	1 ITAG_HUMAN	O75578 homo sapien
11	1030.5	17.2	1151	1 ITAI_HUMAN	P56199 homo sapien
12	1030.5	17.2	1180	1 ITAI_RAT	P18614 rattus norv
13	1012	16.9	1178	1 ITA2_MOUSE	Q62469 mus musculus
14	1007.5	16.8	1170	1 ITA2_BOVIN	P53710 bos taurus
15	1007.5	16.8	1181	1 ITA2_HUMAN	P17301 homo sapien
16	654.5	10.9	1038	1 ITA4_HUMAN	P13612 homo sapien
17	646.5	10.8	1039	1 ITA4_MOUSE	Q00651 mus musculus
18	585.5	9.8	1032	1 ITA4_XENIA	Q91687 xenopus lae
19	584.5	9.8	1035	1 ITA9_HUMAN	Q13797 homo sapien
20	579.5	9.7	1053	1 ITA3_MOUSE	Q62470 mus musculus
21	570.5	9.5	1066	1 ITA3_CRISP	P17852 cricetidae
22	554	9.3	1049	1 ITA5_HUMAN	P08648 homo sapien
23	547	9.1	1039	1 ITAB_HUMAN	P08514 homo sapien
24	540.5	9.0	1066	1 ITA3_HUMAN	P26006 homo sapien
25	538	9.0	1033	1 ITAB_MOUSE	Q9qum0 mus musculus
26	535.5	8.9	1053	1 ITA5_MOUSE	P11688 mus musculus
27	534	8.9	1034	1 ITAV_CHICK	P26008 gallus gall
28	534	8.9	1050	1 ITA5_XENIA	Q06274 xenopus lae
29	535.5	8.8	1396	1 ITA2_DROME	P12080 drosophila
30	497	8.3	1044	1 ITA7_MOUSE	P43406 mus musculus
31	496.5	8.3	1072	1 ITA6_CHICK	P26007 gallus gall
32	493	8.2	1048	1 ITAV_HUMAN	P06756 homo sapien
33	490	8.2	1146	1 ITA1_DROME	Q24247 drosophila

ALIGNMENTS

RESULT 1

ID	ITAD_HUMAN	STANDARD;	PRT;	1162 AA.
AC	Q13349; Q15575; Q15576;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Integrin alpha-D precursor (Leukointegrin alpha D) (CD11d) (ADB2).			
GN	ITGAD.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Splicein;			
RC	MEDLINE=96111956; PubMed=8777714;			
RA	Van der Vliet M., Le Trong H., Wood C.L., Moore P.F., St John T.,			
RA	Staunton D.E., Gallatin W.M.;			
RT	"A novel leukointegrin, alpha d beta 2, binds preferentially to ICAM-			
RT	3.";			
RL	Immunity 3:683-690(1995).			
RN	[2]			
RP	SEQUENCE OF 1-235 FROM N.A.			
RC	MEDLINE=20187620; PubMed=10722744;			
RA	Noti J.D., Johnson A.K., Dillon J.D.;			
RT	"Structural and functional characterization of the leukocyte integrin			
RT	gene CD11d. Essential role of Sp1 and Sp3.";			
RT	J. Biol. Chem. 275:8959-8969(2000).			
RN	[3]			
RP	SEQUENCE OF 467-537; 571-602; 633-666; 788-834 AND 910-1125 FROM N.A.			
RC	MEDLINE=96257236; PubMed=8666289;			
RA	Wong D.A., Davis E.M., LeBeau M., Springer T.A.;			
RT	"Cloning and chromosomal localization of a novel gene-encoding a human			
RT	beta 2-integrin alpha subunit.";			
RL	Gene 171:291-294(1996).			
RN	[4]			
RP	INTERACTION WITH VCAM1.			
RC	MEDLINE=99059842; PubMed=9841932;			
RA	Grayson M.H., Van der Vliet M., Sterbinsky S.A., Michael Gallatin W.,			
RA	Hoffman P.A., Staunton D.E., Bochner B.S.;			
RT	"Alphabeta2 integrin is expressed on human eosinophils and functions			
RT	as an alternative ligand for vascular cell adhesion molecule 1			
RT	(VCAM-1).";			
RT	J. Exp. Med. 188:2187-2191(1998).			
RN	[5]			
RP	INTERACTION WITH VCAM1.			
RC	MEDLINE=99370002; PubMed=10438935;			
RA	Van der Vliet M., Crowe D.T., Hoekstra D., Vazeux R., Hoffman P.A.,			
RA	Grayson M.H., Bochner B.S., Gallatin W.M., Staunton D.E.;			
RT	"The leukocyte integrin alpha D beta 2 binds VCAM-1: evidence for a			
RT	binding interface between I domain and VCAM-1.";			
RL	J. Immunol. 163:1984-1990(1999).			
CC	-I- FUNCTION: INTEGRIN ALPHA-D/BETA-2 IS A RECEPTOR FOR ICAM3 AND			
CC	VCAM1. MAY PLAY A ROLE IN THE ATHEROSCLEROTIC PROCESS SUCH AS			
CC	CLEARING LIPOPROTEINS FROM PLAQUES AND IN PHAGOCYTOSIS OF BLOOD-			

CC Borne pathogens, particulate matter, and senescent erythrocytes
 CC from the blood.
 CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-D
 CC ASSOCIATES WITH BETA-2.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- TISSUE SPECIFICITY: EXPRESSED MODERATELY ON MYELOMONOCYTIC CELL
 CC LINES AND SUBSETS OF PERIPHERAL BLOOD LEUKOCYTES AND STRONGLY ON
 CC TISSUE-SPECIALIZED CELLS, INCLUDING MACROPHAGES FOAM CELLS WITHIN
 CC ATHEROSCLEROTIC PLAQUES, AND ON SPLENIC RED PULP MACROPHAGES.
 CC -1- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
 CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
 CC -1- SIMILARITY: Belongs to the integrin alpha chain family.
 CC -1- SIMILARITY: Contains 1 VWFA domain.
 CC -1- SIMILARITY: Contains 7 FG-GAP repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U37028; AAB38547.1; --
 CC EMBL; U40274; AAB60634.1; --
 CC EMBL; U40275; AAB60635.1; --
 CC EMBL; U40276; AAB60636.1; --
 CC EMBL; U40277; AAB60637.1; --
 CC EMBL; U40279; AAB60638.1; --
 CC EMBL; U40278; AAB60638.1; JOINED.
 CC EMBL; AF187881; AAF62875.1; --
 CC HSSP; P11215; IABX.
 CC Genew; HGNC:6146; ITGAD.
 CC MIM; 602453; --
 CC GO; GO:000305; C:integrin complex; TAS.
 CC GO; GO:004895; F:cell adhesion receptor activity; TAS.
 CC GO; GO:001637; P:cell-cell adhesion; NAS.
 CC GO; GO:0007160; P:cell-matrix adhesion; NAS.
 CC GO; GO:0006955; P:immune response; NAS.
 CC InterPro; IPR00413; Integrin_alpha.
 CC Pfam; PF01839; FG-GAP; 3.
 CC Pfam; PF00357; Integrin_A; 1.
 CC Pfam; PF00092; vwa; 1.
 CC PRINTS; PR01185; INTEGRINA.
 CC PRINTS; PR00453; VWFA DOMAIN.
 CC SMART; SM00191; Int_alpha; 4.
 CC SMART; SM00327; VWA; 1.
 CC PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 CC PROSITE; PS0234; VWFA; 1.
 CC Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 CC Signal; Repeat; Calcium;
 CC Magnesium.
 CC SIGNAL 1 17 POTENTIAL.
 CC CHAIN 18 1162 INTEGRIN ALPHA-D.
 CC DOMAIN 18 1100 EXTRACELLULAR (POTENTIAL).
 CC TRANSMEM 1101 1124 POTENTIAL.
 CC DOMAIN 1125 1162 CYTOPLASMIC (POTENTIAL).
 CC REPEAT 32 85 FG-GAP 1.
 CC REPEAT 2 7 FG-GAP 2.
 CC DOMAIN 150 332 VWFA.
 CC REPEAT 350 400 FG-GAP 3.
 CC REPEAT 401 452 FG-GAP 4.
 CC REPEAT 454 516 FG-GAP 5.
 CC REPEAT 518 576 FG-GAP 6.
 CC REPEAT 581 633 FG-GAP 7.
 CC CA_BIND 465 473 POTENTIAL.
 CC CA_BIND 530 538 POTENTIAL.
 CC CA_BIND 593 601 POTENTIAL.
 CC SITE 1127 1131 GFFKR MOTIF.
 CC DISULFID 67 74 BY SIMILARITY.
 CC DISULFID 106 124 BY SIMILARITY.
 CC DISULFID 655 710 BY SIMILARITY.

FT DISULFID 769 775 BY SIMILARITY.
 FT DISULFID 846 861 BY SIMILARITY.
 FT DISULFID 994 1018 BY SIMILARITY.
 FT DISULFID 1023 1028 BY SIMILARITY.
 FT CARBOHYD 59 57 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 99 99 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 391 391 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 691 691 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 733 733 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 873 873 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 957 957 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1046 1046 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CONFLICT 515 518 MISSING (IN REF. 2).
 FT CONFLICT 825 825 GHPW -> ATP (IN REF. 2).
 FT CONFLICT 984 984 L -> V (IN REF. 2).
 SQ SEQUENCE 1162 AA; 126885 MW; F296A1A35455D77D CRC64;
 Query Match 100.0%; Score 5987; DB 1; Length 1162;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TFGTVLLSVLASYHGFLDVEPTTFQEDAGGFGQSVVQFGGSRVWVGAPLEVAANQT 60
 DB 2 TFGTVLLSVLASYHGFLDVEPTTFQEDAGGFGQSVVQFGGSRVWVGAPLEVAANQT 61
 QY 61 GRLYDCAATGMCQPIPLHTRPAVNMISGLTLAATNGSRLLACGPTLHRVCGENSYSK 120
 DB 62 GRLYDCAATGMCQPIPLHTRPAVNMISGLTLAATNGSRLLACGPTLHRVCGENSYSK 121
 QY 121 GSCLLGSRWEIITQVDPATPECPHOEMDIVFLIDGSGSIDQNDQNMKGFOVAVMGQFE 180
 DB 122 GSCLLGSRWEIITQVDPATPECPHOEMDIVFLIDGSGSIDQNDQNMKGFOVAVMGQFE 181
 QY 181 GTDTLFAIMOYSNLLKTHFTFTFRSPSQOQSLVDPVQLKGLTFTATGTLTVVTLQFHH 240
 DB 182 GTDTLFAIMOYSNLLKTHFTFTFRSPSQOQSLVDPVQLKGLTFTATGTLTVVTLQFHH 241
 QY 241 KNGARKSAKILIVITDQKYPLESDVIPAOKAGIIRYVGVHAFQGTARQELN 300
 DB 242 KNGARKSAKILIVITDQKYPLESDVIPAOKAGIIRYVGVHAFQGTARQELN 301
 QY 301 TISSAPPQDHVFKVDNFAALGSIQKQKEKIYAVEGTQSRASSSFQHEMSQEGFSTALTM 360
 DB 302 TISSAPPQDHVFKVDNFAALGSIQKQKEKIYAVEGTQSRASSSFQHEMSQEGFSTALTM 361
 QY 361 DGLFLGAVGSGGSAFLYPNNMPTFINMSQENVMDRDSYLGYSTELALWKGQVNLVL 420
 DB 362 DGLFLGAVGSGGSAFLYPNNMPTFINMSQENVMDRDSYLGYSTELALWKGQVNLVL 421
 QY 421 GAPRYQHTGKAVIPTQVSRQWRKKAETGTQIGSYFGASLCSVDVDSGDSTDLILIGAPH 480
 DB 422 GAPRYQHTGKAVIPTQVSRQWRKKAETGTQIGSYFGASLCSVDVDSGDSTDLILIGAPH 481
 QY 481 YYEOTRGQSVCLPRGQVQWOCDAVLRGEGHPWGRFGAALTVLGDVNEKLIIDVAI 540
 DB 482 YYEOTRGQSVCLPRGQVQWOCDAVLRGEGHPWGRFGAALTVLGDVNEKLIIDVAI 541
 QY 541 GAPGEQENRGAVILFHGASBSGISPSHSQRIASSQLSPRLQYFGQALSGGQDLTQDGLMD 600
 DB 542 GAPGEQENRGAVILFHGASBSGISPSHSQRIASSQLSPRLQYFGQALSGGQDLTQDGLMD 601
 QY 601 LAVGARGQVLLRSLPVLKGVAMRFPSEVAVKAVYRCWEKPSALBAGATVCLTIQKS 660
 DB 602 LAVGARGQVLLRSLPVLKGVAMRFPSEVAVKAVYRCWEKPSALBAGATVCLTIQKS 661
 QY 661 SLDLQDIQSSVRFDLALDPGLTSLRAIFNETKNPTLTRKTLGLGHICHTLKLPLDCV 720
 DB 662 SLDLQDIQSSVRFDLALDPGLTSLRAIFNETKNPTLTRKTLGLGHICHTLKLPLDCV 721
 QY 721 EDVYSPILHLNFSLVREPISPNLPRVLAVGQDLFTASLPPEKNCQGGLCEGLGV 780

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2004, 09:36:23 ; Search time 56 Seconds
(without alignments)
6541.368 Million cell updates/sec

Title: US-09-891-943-2

Perfect score: 5987
Sequence: 1 TFGTVLLSVLASHYHGFLND.....DTATSGDDFSCVAPNVLIS 1161

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 segs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp Vertebrate.*
- 14: sp Unclassified.*
- 15: sp_virus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4415.5	73.8	1161	11 Q9QVE7	Q9QVE7 rattus norv
2	3824	63.9	1169	4 Q8IVA6	Q8IVA6 homo sapien
3	3581.5	59.8	1169	11 Q9QXH4	Q9QXH4 mus musc
4	3314	55.4	1151	11 Q9J130	Q9J130 rattus norv
5	2829	47.3	920	6 Q8Z984	Q8Z984 sus scrofa
6	2755	46.0	1036	11 Q8CAV3	Q8CAV3 mus musc
7	1451.5	24.2	1161	11 Q9WTV4	Q9WTV4 mus musc
8	1447	24.2	1160	11 Q9R200	Q9R200 mus musc
9	1318	22.0	1086	4 Q96HB1	Q96HB1 homo sapien
10	1297.5	21.7	1196	13 Q98TF1	Q98TF1 cyprinus ca
11	1263	21.1	1187	13 Q98TF0	Q98TF0 cyprinus ca
12	1220.5	20.4	927	6 Q8HZV0	Q8HZV0 bos taurus
13	1164.5	19.5	1167	11 Q88340	Q88340 rattus norv
14	1138.5	19.0	1167	11 Q88341	Q88341 rattus norv
15	1096.5	18.3	1188	11 Q7TQC3	Q7TQC3 mus musc
16	1041	17.4	1171	13 Q42094	Q42094 gallus gall

17	1017	17.0	348	4 Q8TBS5	Q8TBS5 homo sapien
18	1015.5	17.0	1038	11 Q8BS01	Q8BS01 mus musc
19	984.5	16.4	895	11 Q9WUF8	Q9WUF8 mus sp. itg
20	972.5	16.2	1160	6 Q8MKF4	Q8MKF4 felis silve
21	812	13.6	1332	5 Q9BPQ8	Q9BPQ8 halocynthia
22	771.5	12.9	823	4 Q8WY18	Q8WY18 homo sapien
23	722	12.1	823	11 Q8C584	Q8C584 mus musc
24	656.5	11.0	1032	11 Q61989	Q61989 mus musc
25	643	10.7	780	13 Q9U671	Q9U671 xenopus lae
26	638	10.7	1054	5 Q9U6S1	Q9U6S1 strongyloce
27	629.5	10.5	1474	5 Q8G687	Q8G687 pseudoplusi
28	629	10.5	205	11 Q63001	Q63001 rattus norv
29	622	10.4	1033	6 Q9BGU3	Q9BGU3 bos taurus
30	621.5	10.4	1036	11 Q91YD5	Q91YD5 mus musc
31	620	10.4	1034	13 Q98TT7	Q98TT7 gallus gall
32	615	10.3	1041	5 Q9UB90	Q9UB90 lytechinus
33	611	10.2	1041	5 Q76378	Q76378 lytechinus
34	610	10.2	257	11 Q8C270	Q8C270 mus musc
35	568.5	9.5	1036	6 Q7YRP8	Q7YRP8 equus cabal
36	548	9.2	1034	6 Q9TUN4	Q9TUN4 cryctolagus
37	537	9.0	1016	13 Q91779	Q91779 xenopus lae
38	532.5	8.9	1053	11 Q80VP5	Q80VP5 mus musc
39	526	8.8	1119	5 Q8G688	Q8G688 pseudoplusi
40	509	8.5	231	4 Q8N882	Q8N882 homo sapien
41	502.5	8.4	1012	11 Q70304	Q70304 mus musc
42	490	8.2	1007	6 Q9GK48	Q9GK48 bos taurus
43	490	8.2	1033	13 Q42598	Q42598 xenopus lae
44	490	8.2	1049	5 Q8SY51	Q8SY51 drosophila
45	488	8.2	1047	6 Q9MZD6	Q9MZD6 bos taurus

ALIGNMENTS

RESULT 1

Q9QVE7	Q9QVE7	PRELIMINARY;	PRT; 1161 AA.
AC	Q9QVE7;		
DT	01-MAY-2000 (TREMBLrel. 13, Created)		
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)		
DT	01-JUN-2003 (TREMBLrel. 24, Last annotation update)		
DE	Alpha D integrin.		
OS	Rattus norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
OX	NCBI_TaxID=10116;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Sprague-Dawley;		
RA	O'Brien M.W., Vandervieren M., Kilgannon P.D., Dietsch G.,		
RA	Gallatin W.M.;		
RT	"Cloning of rat alpha D, a novel beta 2 integrin.";		
RL	Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AF021334; AAF21241.1; -		
DR	HSSP; P11215; 1BHQ.		
DR	GO; GO:0008305; C:integrin complex; IEA.		
DR	GO; GO:0004895; P:cell adhesion receptor activity; IEA.		
DR	GO; GO:0007160; P:cell-matrix adhesion; IEA.		
DR	GO; GO:0007229; P:integrin-mediated signaling pathway; IEA.		
DR	InterPro; IPR000413; Integrin_alpha.		
DR	Pfam; PF01839; FG-GAP; 3.		
DR	Pfam; PF00357; Integrin_A; 1.		
DR	Pfam; PF00092; vwa; 1.		
DR	PRINTS; PR01185; INTEGRINA.		
DR	PRINTS; PR00453; VWFADOMAIN.		
DR	SMART; SMC0191; Int_alpha; 4.		
DR	SMART; SMC0327; VWA; 1.		
DR	PROSITE; PS00242; INTEGRIN_ALPHA; 1.		
DR	PROSITE; PS50234; VWFA; 1.		
KW	Integrin.		
SQ	SEQUENCE 1161 AA; 126600 MW; 2258491A984N705E CFC64;		

```
Query Match      73.8%; Score 4415.5; DB 11; Length 1161;
Best Local Similarity 73.8%; Pred. No. 0;
Matches 853; Conservative 111; Mismatches 187; Indels 5; Gaps 4;

QY 3 GTVLLLS--VLASVHGFNLVDEPTTFOEDAGGFGQSVVQFGGSRVWVGAPELVAAQT 60
DB 4 GVILLGCVLASCHGSLNLDVEPIVFRDAAAFSGQTVVQFGGSRVWVGAPELVAAQT 63

QY 61 GRLYDCAATGMCQPIPLHTRPAVNMVSLGLTLAASTNGSRLLACGPTLHRVCGENSYSK 120
DB 64 GRLYDCAATGMCQPIVLRSPLEAVNMVSLGLSVTTATNNAQLLACGPTAQRACVKNMYAK 123

QY 121 GSCLLGSRWEIIQTVPDAPTECPHQMEDIIVFLIDGSGSIDQNDFNQMKGFQVAVMGQPE 180
DB 124 GSCLLGSSIQFOAVPASMECPQEQMDIAFLIDGSGSINQDFQAKDFYKALMGFEA 183

QY 181 GTDTLFAIMOYSLNLLKHFTFTQPRSPSOQSLVDPIVQLKGLTFTATGILTVVTLFHH 240
DB 184 STSTLFSIMOYSLNLLKHFTFTQPRSPSOQSLVDPIVQLKGLTFTATGILTVVTLFHH 243

QY 241 KNGARKSAKILLIVITDQKYKDPLEYSVDIPOAEKAGIIRYAIYGVGHAFQGPRTARQELN 300
DB 244 KNGSRSAKILLIVITDQKYKDPLEYSVDIPOAEKAGIIRYAIYGVGHAFQGPRTARQELN 303

QY 301 TISSAPPQDHVFKVDNPAALGSIQKQLEKIYAVEGTQSRASSSFQHEMSQEGFSTALTM 360
DB 304 TISSAPPQDHVFKVDNPAALGSIQKQLEKIYAVEGTQSRASSSFQHEMSQEGFSTALTM 363

QY 361 DGLFLGAVGFSWGGAFVPPNMSPTFNNSQENVDMDSDYLGSTELANNGVONLV 420
DB 364 DGPVLGAVGFSWGGAFVPPNMSPTFNNSQENVDMDSDYLGSTELANNGVONLV 423

QY 421 GAPRYQHTKAVIPTQVSRQWRKAEYTGTOIGSYFGASLCSVDVSDGSDTLILGAPH 480
DB 424 GAPRHQHTKAVIPTQVSRQWRKAEYTGTOIGSYFGASLCSVDVSDGSDTLILGAPH 483

QY 481 YYBQTRGGQVSVCEPLPGQVQVOCDAVLREGEQHPWRGGAALTVLGDVNEKLDVAI 540
DB 484 YYBQTRGGQVSVCEPLPGQVQVOCDAVLREGEQHPWRGGAALTVLGDVNEKLDVAI 542

QY 541 GAPQEQNRGAVLIFHGASRSGISPSHSQRIASSQSLSPRIQYFGQALSGQDITDGLMD 600
DB 543 GAPQEQNRGAVLIFHGASRSGISPSHSQRIASSQSLSPRIQYFGQALSGQDITDGLMD 602

QY 601 LAVGARGQVLLSLPLVKVGVAMRSPVEVAVNVRCHKEKPSALBAGDAITVCLTIQKS 660
DB 603 LAVGARGQVLLSLPLVKVGVAMRSPVEVAVNVRCHKEKPSALBAGDAITVCLTIQKS 662

QY 661 SLDDQIGIQSVRFDLALDPLGRLTSRAIFNETKNTPLTRKTLGLGHCETLKLPLDCV 720
DB 663 SPDLLGNVQSVAYDALDPLGRLTSRAIFNETKNTPLTRKTLGLGHCETLKLPLDCV 722

QY 721 EDVVSPIILHNLPSLVREPIPSQNLRLPVLAVGSDQLFTASLPFRNCGQDGLCEGDLGV 780
DB 723 EDVVSPIILHNLPSLVREPIPSQNLRLPVLAVGSDQLFTASLPFRNCGQDGLCEGDLGV 781

QY 781 TLSPSGIQTILTVGSSLELNVITVWNAQSDSYGVVSVLYPAGLSHRRYSVGAQKQPHQSA 840
DB 782 SFNPSGLQVLVVGSGPELTVVWNAQSDSYGVVSVLYPAGLSHRRYSVGAQKQPHQSA 840

QY 841 LRLACEVTPTEDEGLRSRCSVNHPIPHFEGSGNTPVITFDVSYKATLGDRLMLRASASSE 900
DB 841 LRLACEVTPTEDEGLRSRCSVNHPIPHFEGSGNTPVITFDVSYKATLGDRLMLRASASSE 900

QY 901 NKKASSKATFQLEPLVKYAVYTMISQEBSTKYENFATSDKKMKEAEHRYVNNLSQR 960
DB 901 NKKASSKATFQLEPLVKYAVYTMISQEBSTKYENFATSDKKMKEAEHRYVNNLSQR 960

QY 961 DLAISINFWPVLNGVAVDVNVAEPQSLPCVSRKPPQHSDFPTQISRSPLMDCSIA 1020
DB 961 KLVAVNFWPVLNGVAVDVNVAEPQSLPCVSRKPPQHSDFPTQISRSPLMDCSIA 1020

QY 1021 DCUQRCDCVPSFQBELDFTLKNLSFGVWRETLQKVLVVSVAITFTDTSVYSQLPQG 1080
```

```
DB 1021 DCLHFRCDIPSLDQDELDFTLKNLSFGVWRETLQKVLVVSVAITFTDTSVYSQLPQG 1080

QY 1081 EAFRAQEMVLEDEYVNAIPIMSGSVGALLLALITATLYKLGFKRHYKEMLEDKP 1140
DB 1081 EAFRAQEMVLEDEYVNAIPIMSGSVGALLLALITATLYKLGFKRHYKEMLEDKP 1140

QY 1141 EDTATFSGDDDFSCVAP 1156
DB 1141 ADPVTAGQADFGCETP 1156

RESULT 2
Q81VA6
ID Q81VA6 PRELIMINARY; PRT; 1169 AA.
AC Q81VA6:
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Integrin, alpha X (Antigen CD11C (p150), alpha polypeptide).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_taxid=9606;
RN [1]
SEQUENCE FROM N.A.
RP TISSUE=Blood;
RC Strausberg R.;
RA Submitted (SSP-2002) to the EMBL/GenBank/DBSJ databases.
DR EMBL; BC038227; AAH38237.1; -.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; P:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR02035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS0234; VFWA; 1.
DR PROSITE; PS0234; VFWA; 1.
SQ SEQUENCE 1169 AA; 128521 MW; A178484FEFC79EB6 CRC64;

Query Match      63.9%; Score 3824; DB 4; Length 1169;
Best Local Similarity 65.8%; Pred. No. 4.6e-285;
Matches 758; Conservative 135; Mismatches 253; Indels 6; Gaps 5;

QY 1 TFGTVLLSLVSLVHGFNLVDEPTTFOEDAGGFGQSVVQFGGSRVWVGAPELVAAQT 60
DB 4 TRALLLFTALATSLGFLNLDTEELTAFVDSAGFSDSVVQVANSVWVGAPELVAAQT 63

QY 61 GRLYDCAATGMCQPIPLHTRPAVNMVSLGLTLAASTNGSRLLACGPTLHRVCGENSYSK 120
DB 64 GRLYDCAATGMCQPIPLHTRPAVNMVSLGLTLAASTNGSRLLACGPTLHRVCGENSYSK 123

QY 121 GSCLLGSRWEIIQTVPDAPTECPHQMEDIIVFLIDGSGSIDQNDFNQMKGFQVAVMGQPE 180
DB 124 GSCLLGSRWEIIQTVPDAPTECPHQMEDIIVFLIDGSGSIDQNDFNQMKGFQVAVMGQPE 182

QY 181 GTDTLFAIMOYSLNLLKHFTFTQPRSPSOQSLVDPIVQLKGLTFTATGILTVVTLFHH 240
DB 183 RPSQFSLMQFSNKFQTHFTFEFRSSNPLSLASVHQLQGFYTTATAIQNVVHRLFA 242

QY 241 KNGARKSAKILLIVITDQKYKDPLEYSVDIPOAEKAGIIRYAIYGVGHAFQGPRTARQELN 300
DB 243 SYGARDAAKILLIVITDQKYKDPLEYSVDIPOAEKAGIIRYAIYGVGHAFQGPRTARQELN 302

QY 301 TISSAPPQDHVFKVDNPAALGSIQKQLEKIYAVEGTQSRASSSFQHEMSQEGFSTALTM 360
DB 303 DIASKPSQEHIFKVEDFDALDKIQNLKELIFALTEGTTTSSSSSFELEMAQEGSAVFTP 362
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2004, 09:32:28 ; Search time 65 seconds
(without alignments)
5046.733 Million cell updates/sec

Title: US-09-891-943-2
Perfect score: 5987
Sequence: 1 TFGTVLLSVLASVHGFLND.....DTATFGDGFSCVAPNVPLS 1161

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues
Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1980s:*
3: geneseqp2000s:*
4: geneseqp2000s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5987	100.0	1161	2 AAR78166	Aar78166 Human bet
2	5987	100.0	1161	2 AAW23049	Aaw23049 Human bet
3	5987	100.0	1161	2 AAW57491	Aaw57491 Human bet
4	5987	100.0	1161	2 AAW72825	Aaw72825 Human alp
5	5987	100.0	1161	3 AAB07359	Aab07359 Human alp
6	5987	100.0	1161	5 ABG61468	Abg61468 Human bet
7	5981	99.9	1161	2 AAW65089	Aaw65089 Human bet
8	5981	99.9	1161	2 AAW73342	Aaw73342 Human alp
9	5971.5	99.7	1161	2 AAW23064	Aaw23064 Human bet
10	5971.5	99.7	1161	2 AAW65106	Aaw65106 Human bet
11	5971.5	99.7	1161	2 AAW72837	Aaw72837 Human alp
12	5971.5	99.7	1161	2 AAW73343	Aaw73343 Human alp
13	5971.5	99.7	1161	3 AAB07376	Aab07376 Human alp
14	5971.5	99.7	1161	5 ABG61485	Abg61485 Human bet
15	4403.5	73.6	1161	2 AAW23062	Aaw23062 Rat beta
16	4403.5	73.6	1161	2 AAW60004	Aaw60004 Rat alpha
17	4403.5	73.6	1161	2 AAW72824	Aaw72824 Rat alpha
18	4403.5	73.6	1161	3 AAB07374	Aab07374 Rat alpha
19	4403.5	73.6	1161	5 ABG61483	Abg61483 Rat beta2
20	4397.5	73.5	1161	2 AAW65104	Aaw65104 Rat beta-
21	4397.5	73.5	1161	2 AAW73345	Aaw73345 Rat alpha
22	4395.5	73.4	1161	2 AAR78169	Aar78169 Rat alpha
23	4365.5	72.9	1151	2 AAR78179	Aar78179 Rat beta
24	4364.5	72.9	1151	2 AAW23059	Aaw23059 Rat beta
25	4364.5	72.9	1151	2 AAW60001	Aaw60001 Rat alpha

26	4364.5	72.9	1151	2 AAW65101	Aaw65101 Rat beta-
27	4364.5	72.9	1151	2 AAW72834	Aaw72834 Rat alpha
28	4364.5	72.9	1151	2 AAW73344	Aaw73344 Rat alpha
29	4364.5	72.9	1151	3 AAB07371	Aab07371 Rat alpha
30	4364.5	72.9	1151	5 ABG61480	Abg61480 Rat beta2
31	4305.5	71.9	1161	2 AAW23061	Aaw23061 Mouse bet
32	4305.5	71.9	1161	2 AAW60003	Aaw60003 Mouse alp
33	4305.5	71.9	1161	2 AAW65103	Aaw65103 Mouse bet
34	4305.5	71.9	1161	2 AAW72836	Aaw72836 Mouse alp
35	4305.5	71.9	1161	2 AAW73347	Aaw73347 Mouse alp
36	4305.5	71.9	1161	3 AAB07373	Aab07373 Mouse alp
37	4305.5	71.9	1161	5 ABG61482	Abg61482 Mouse bet
38	4303.5	71.9	1161	2 AAR78168	Aar78168 Mouse alp
39	4278.5	71.5	1155	2 AAR78167	Aar78167 Mouse alp
40	4278.5	71.5	1155	2 AAW23060	Aaw23060 Mouse bet
41	4278.5	71.5	1155	2 AAW60002	Aaw60002 Mouse alp
42	4278.5	71.5	1155	2 AAW65102	Aaw65102 Mouse bet
43	4278.5	71.5	1155	2 AAW72835	Aaw72835 Mouse alp
44	4278.5	71.5	1155	2 AAW73346	Aaw73346 Mouse alp
45	4278.5	71.5	1155	3 AAB07372	Aab07372 Mouse alp

ALIGNMENTS

RESULT 1
AAR78166
ID AAR78166 standard; protein; 1161 AA.
XX

AC AAR78166;

XX 28-DEC-1995 (first entry)

XX Human beta-2 integrin alpha-d.

XX Beta-2 integrin alpha-d subunit; antinflammatory; arteriosclerosis;
KW inflammatory bowel disease; asthma.

XX Homo sapiens.

XX Location/Qualifiers

FT Key
FT Domain
FT Region

FT Binding-site
FT Binding-site
FT Binding-site

FT Region
FT Domain
FT Domain

FT Binding-site
FT Binding-site
FT Binding-site

FT Region
FT Domain
FT Domain

FT Binding-site
FT Binding-site
FT Binding-site

FT Region
FT Domain
FT Domain

FT Binding-site
FT Binding-site
FT Binding-site

FT Region
FT Domain
FT Domain

FT Binding-site
FT Binding-site
FT Binding-site

FT Region
FT Domain
FT Domain

FT Binding-site
FT Binding-site
FT Binding-site

FT Region
FT Domain
FT Domain

FT Binding-site
FT Binding-site
FT Binding-site

FT Region
FT Domain
FT Domain

FT Binding-site
FT Binding-site
FT Binding-site

FT Region
FT Domain
FT Domain

FT Binding-site
FT Binding-site
FT Binding-site

FT Region
FT Domain
FT Domain

FT Binding-site
FT Binding-site
FT Binding-site

XX Alpha sub-unit polypeptide of human beta 2 integrin - used to identify
PT potential antiinflammatory agents, for the treatment of graft
PT arteriosclerosis, inflammatory bowel disease, asthma, etc.
XX
PS
XX Claim 7; Page 82-87; 172pp; English.
XX
CC A probe based on a partial cDNA clone (given in AAQ91727) of canine alpha
CC -TM1 was used to screen a human spleen cDNA library to identify clone
CC 19A2 encoding the beta-1 integrin alpha-subunit, alpha-d. The cDNA was
CC manipulated to allow expression of recombinant alpha-d subunit in COS and
CC CHO cells
XX
XX Sequence 1161 AA;
SQ
Query Match 100.0%; Score 5987; DB 2; Length 1161;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 TFGTVLLSVLASHYHGNLDVEPTTFQEDAGFGQSVVQFGSRLVWGAPLEVVAAQT 60
DB 1 TFGTVLLSVLASHYHGNLDVEPTTFQEDAGFGQSVVQFGSRLVWGAPLEVVAAQT 60
OY 61 GRLYDCAAATGMCQPTPLHIREAVNMSLGLTLASTNGSRLLAGCPTLHRCVGENSYK 120
DB 61 GRLYDCAAATGMCQPTPLHIREAVNMSLGLTLASTNGSRLLAGCPTLHRCVGENSYK 120
OY 121 GSCLLGSRWEIITQVPDAPTECFHOEMDIFVLIDSGSIDQNDNOMKGFVQAVMGQFE 180
DB 121 GSCLLGSRWEIITQVPDAPTECFHOEMDIFVLIDSGSIDQNDNOMKGFVQAVMGQFE 180
OY 181 GTDTLFAIMQYSLNKLHHTFTFRSPSQOQLVDPIVQLKGLTPTATGILTVTLQFLFH 240
DB 181 GTDTLFAIMQYSLNKLHHTFTFRSPSQOQLVDPIVQLKGLTPTATGILTVTLQFLFH 240
OY 241 KNGARKSAXKILVITDQKXKDPLEYSVPIPOEXAGIIRYAIGVGHAFQGPTRQELN 300
DB 241 KNGARKSAXKILVITDQKXKDPLEYSVPIPOEXAGIIRYAIGVGHAFQGPTRQELN 300
OY 301 TISSAPQDHVFKVDNFALGSKLOKQIYAVEGTQSRASSFOHENSQSGFSTALTM 360
DB 301 TISSAPQDHVFKVDNFALGSKLOKQIYAVEGTQSRASSFOHENSQSGFSTALTM 360
OY 361 DGLFLGAVGSFSGGAFLYPPNMTFFINMSQENVDMRDSYLGXSTELALWKGVQNLVL 420
DB 361 DGLFLGAVGSFSGGAFLYPPNMTFFINMSQENVDMRDSYLGXSTELALWKGVQNLVL 420
OY 421 GAPRYOHTGKAVITQVSRQWRKKAETGTQIGSYFGASLCSVDVDSGSDTLILIGAPH 480
DB 421 GAPRYOHTGKAVITQVSRQWRKKAETGTQIGSYFGASLCSVDVDSGSDTLILIGAPH 480
OY 481 YFEQTRGGQVSVCPLEPRGQVQCCDAVLGRGQHPWGRFGAALTVLGDVNEKLDIVAI 540
DB 481 YFEQTRGGQVSVCPLEPRGQVQCCDAVLGRGQHPWGRFGAALTVLGDVNEKLDIVAI 540
OY 541 GAPCEQNRGAVLPHGASBSGSPSHORSQIASLSQSPRLQYFGQALSGQDLTQDGLMD 600
DB 541 GAPCEQNRGAVLPHGASBSGSPSHORSQIASLSQSPRLQYFGQALSGQDLTQDGLMD 600
OY 601 LAVGARGQVLLRLSLPVLKVGVMRSPVEVAKAVYRCWEKPSALEAGDATVCLTIQKS 660
DB 601 LAVGARGQVLLRLSLPVLKVGVMRSPVEVAKAVYRCWEKPSALEAGDATVCLTIQKS 660
OY 661 SLDQLGDIQSVRFDLALDPRLTSTRAIFNETKNPTLTRKTLGLGHIHCEETKLLLPDCV 720
DB 661 SLDQLGDIQSVRFDLALDPRLTSTRAIFNETKNPTLTRKTLGLGHIHCEETKLLLPDCV 720
OY 721 EDVVSPIILHLNLSLVREPIPSQNLRLPVLAVGSQDLFTASLPFEKNCQDGLCEGDLGV 780
DB 721 EDVVSPIILHLNLSLVREPIPSQNLRLPVLAVGSQDLFTASLPFEKNCQDGLCEGDLGV 780
OY 781 TLSFSGQLQTLTVGSSLELNVITVMNAGEDSVGTVVSLLYYPAGLSHRRVSGAQKQPHQSA 840

DB 781 TLSFSGQLQTLTVGSSLELNVITVMNAGEDSVGTVVSLLYYPAGLSHRRVSGAQKQPHQSA 840
OY 841 LRLACETVPTDEGLRSSRCNVNHPHFHEGSGNGTFIVTFDVSYKATLGDRLMLRASASSE 900
DB 841 LRLACETVPTDEGLRSSRCNVNHPHFHEGSGNGTFIVTFDVSYKATLGDRLMLRASASSE 900
OY 901 NNKASSKATFQLELPVKYAVYTMISROBESTKYFNFAISDEKMKKEAHRVNNLSOR 960
DB 901 NNKASSKATFQLELPVKYAVYTMISROBESTKYFNFAISDEKMKKEAHRVNNLSOR 960
OY 961 DLAISINFWPVLINGVAVWVDVMEAPSOGLPCVSRKPKQHSDFLTQISRSPLMDCSIA 1020
DB 961 DLAISINFWPVLINGVAVWVDVMEAPSOGLPCVSRKPKQHSDFLTQISRSPLMDCSIA 1020
OY 1021 DCQFRCDVPSFVQBELDFTLKNLSFGWVRETLOKKVLVWSVAEITDTSVYSQLPQG 1080
DB 1021 DCQFRCDVPSFVQBELDFTLKNLSFGWVRETLOKKVLVWSVAEITDTSVYSQLPQG 1080
OY 1081 EAFMRQEMWLEDEEVYNAIPIIMGSSVGCALLLALITATLYKLGFFKRYKEMLEDKP 1140
DB 1081 EAFMRQEMWLEDEEVYNAIPIIMGSSVGCALLLALITATLYKLGFFKRYKEMLEDKP 1140
OY 1141 EDTATFSGDDFSCVAPNVPLS 1161
DB 1141 EDTATFSGDDFSCVAPNVPLS 1161
RESULT 2
AAW23049
ID AAW23049 standard; protein; 1161 AA.
XX
AC AAW23049;
DT 24-FEB-1998 (first entry)
XX
DE Human beta 2 integrin alpha d subunit.
XX
KW Beta 2 integrin alpha d subunit; human; cell migration; cell adhesion;
KW phagocytosis; diabetes; atherosclerosis; multiple sclerosis; asthma;
KW psoriasis; lung inflammation; acute respiratory distress syndrome;
KW rheumatoid arthritis.
XX
OS Homo sapiens.
XX
FH Key
FT Domain 17. .1108
FT Domain /label= Extracellular_domain
FT Domain 150. .352
FT Domain /note= "region homologous to the I (insertion) domain
FT Domain 1109. .1128
FT Domain /label= Transmembrane_domain
FT Domain /note= "homologous to the human CD11c transmembrane
FT Domain 1129. .1161
FT Domain /label= Cytoplasmic_domain
XX
PN WO9731099-A1.
XX
PD 28-AUG-1997.
XX
PF 24-FEB-1997; 97WO-US002713.
XX
PR 22-FEB-1996; 96US-00605672.
XX
PA (ICOS-) ICOS CORP.
XX
PI Gallatin WW, Van Der Vieren M;
XX
DR WPI; 1997-435154/40.
DR N-PSDB; AAT73220.
XX
PT Hyridoma 199M and antibody secreted by it - specific for new rat beta2

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2004, 09:40:59 ; Search time 54 Seconds
(without alignments)
6048.782 Million cell updates/sec

Title: US-09-891-943-2
Perfect score: 5987
Sequence: 1 TFGTVLLSVLASHYHGFNLD.....DTATFGSDDFSCVAPNVPLS 1161

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1155919 seqs, 281338677 residues

Total number of hits satisfying chosen parameters: 1155919

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

ALIGNMENTS

RESULT 1
US-09-350-259-2
; Sequence 2, Application US/09350259
; Patent No. US20020062008A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US20020062008A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/350,259
; CURRENT FILING DATE: 1999-07-08
; EARLIER APPLICATION NUMBER: 09/193,043
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173,497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 05/943,363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1161
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-350-259-2

Query Match 100.0%; Score 5987; DB 9; Length 1161;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TFGTVLLSVLASHYHGFNLDVEPTIFQDAGFGQSVVQFGSRLVVGAPLEVAANT 60
Db 1 TFGTVLLSVLASHYHGFNLDVEPTIFQDAGFGQSVVQFGSRLVVGAPLEVAANT 60
Qy 61 GRLYDCAATGMCP:PLHRIPEAVNMS:GLTLAASTNGSRLLAGCPTLHRVCGENSYSK 120

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5987	100.0	1161	9	US-09-350-259-2
2	5987	100.0	1161	10	US-09-891-943-2
3	5971.5	99.7	1161	9	US-09-350-259-59
4	5971.5	99.7	1161	10	US-09-891-943-59
5	4403.5	73.6	1161	9	US-09-350-259-55
6	4403.5	73.6	1161	10	US-09-891-943-55
7	4364.5	72.9	1151	9	US-09-350-259-37
8	4364.5	72.9	1151	10	US-09-891-943-37
9	4305.5	71.9	1161	9	US-09-350-259-53
10	4305.5	71.9	1161	10	US-09-891-943-53
11	4278.5	71.5	1155	9	US-09-350-259-46
12	4278.5	71.5	1155	10	US-09-891-943-46
13	3815	62.6	1163	15	US-10-116-275-204
14	3746	62.6	1163	9	US-09-350-259-4
15	3746	62.6	1163	10	US-09-891-943-4

Db 61 GRLYDCAATGMCQIPPHIRPEAVNMSLGLTLASTNGSRLLACGPTLHRVCGENSYSK 120
Qy 121 GSCLLLSRWELIQTVPDAPTECPHQMNDIVFLIDSGSIDQDNFNQMGFVQAVMGQFE 180
Db 121 GSCLLLSRWELIQTVPDAPTECPHQMNDIVFLIDSGSIDQDNFNQMGFVQAVMGQFE 180
Qy 181 GDTTLFALMOYNSLLKIHFTFTQFTSPSQSLVDPPIVOLKGLTPTATGILTVVTLFHH 240
Db 181 GDTTLFALMOYNSLLKIHFTFTQFTSPSQSLVDPPIVOLKGLTPTATGILTVVTLFHH 240
Qy 241 KNGARKSAKILIVITDGQKIDPLEYSDVIPQAEKAGIIRYAIGVGHAFQGPPTARQELN 300
Db 241 KNGARKSAKILIVITDGQKIDPLEYSDVIPQAEKAGIIRYAIGVGHAFQGPPTARQELN 300
Qy 301 TISSAPPQDHVKVDNFAALGSIQKQLEKIYAVGTSRASSSQHESQEGSFSTALTM 360
Db 301 TISSAPPQDHVKVDNFAALGSIQKQLEKIYAVGTSRASSSQHESQEGSFSTALTM 360
Qy 361 DGLFLGAVGSFWSGGAFLYPPNMSPTFINMSQENVMDRDSYLGYSTELALWKGVQNLVL 420
Db 361 DGLFLGAVGSFWSGGAFLYPPNMSPTFINMSQENVMDRDSYLGYSTELALWKGVQNLVL 420
Qy 421 GAPRYOHTGKAVITQVSRQWRKKAETVGTQIGSYFGASLCSVDVDSGSDTLILIGAPH 480
Db 421 GAPRYOHTGKAVITQVSRQWRKKAETVGTQIGSYFGASLCSVDVDSGSDTLILIGAPH 480
Qy 481 YTEQTRGGQVSVCPFRGQVQWQCDVLRGEGHPWGRFGAALTVLGDVNEKLDIVAI 540
Db 481 YTEQTRGGQVSVCPFRGQVQWQCDVLRGEGHPWGRFGAALTVLGDVNEKLDIVAI 540
Qy 541 GAPGEQENRGAVYLFHGAESGISPSHSQRIASSQLSPRLQYFGQALSGGQDLTQDGLMD 600
Db 541 GAPGEQENRGAVYLFHGAESGISPSHSQRIASSQLSPRLQYFGQALSGGQDLTQDGLMD 600
Qy 601 LAVGARGQVLLRSPLVKGVAMRSPVEVAKAVYRCWEKPSALEAGDATVCLTIQKS 660
Db 601 LAVGARGQVLLRSPLVKGVAMRSPVEVAKAVYRCWEKPSALEAGDATVCLTIQKS 660
Qy 661 SLDQLGDIQSSVRFDLALDPRLTSRAIFNETHNPTLIRKTLGIGHCETLKLILPCV 720
Db 661 SLDQLGDIQSSVRFDLALDPRLTSRAIFNETHNPTLIRKTLGIGHCETLKLILPCV 720
Qy 721 EDVSPILHLNFSLVREPIPSQNLRPVLAVGSQDLFTASLPFEKNCQDGLCEGDLGV 780
Db 721 EDVSPILHLNFSLVREPIPSQNLRPVLAVGSQDLFTASLPFEKNCQDGLCEGDLGV 780
Qy 781 TISFSGLOTLTVGSSLELNVITVWVAGSDSYGVVSLYYPAGLSHRRVSGAKQPHQSA 840
Db 781 TISFSGLOTLTVGSSLELNVITVWVAGSDSYGVVSLYYPAGLSHRRVSGAKQPHQSA 840
Qy 841 LRLACETVPTEDGLSRSSCVNHPITFHEGSGNTRIVTFDVSYKATLGRMLMRASASSE 900
Db 841 LRLACETVPTEDGLSRSSCVNHPITFHEGSGNTRIVTFDVSYKATLGRMLMRASASSE 900
Qy 901 NNXASSKATFQLELPVKYAVYTMISQBEESTKYFNFATSDKKMKAEHRYRVNLSOR 960
Db 901 NNXASSKATFQLELPVKYAVYTMISQBEESTKYFNFATSDKKMKAEHRYRVNLSOR 960
Qy 961 DLAISINFWPVLINGVAVMDVWNEAPSQSLPCVSRKPPQSDFTQISRSPMLDCSIA 1020
Db 961 DLAISINFWPVLINGVAVMDVWNEAPSQSLPCVSRKPPQSDFTQISRSPMLDCSIA 1020
Qy 1021 DCLQFRCDVPSFVSQBELDFTLXGNLSFGWVRETLQKVLVWSVABITFDTSVYSQLPGQ 1080
Db 1021 DCLQFRCDVPSFVSQBELDFTLXGNLSFGWVRETLQKVLVWSVABITFDTSVYSQLPGQ 1080
Qy 1081 EAFMRAQEMVLEDEYVNAIPIIMSSVGCALILALITATLYKLGFPKRRHYKEMLEDKP 1140
Db 1081 EAFMRAQEMVLEDEYVNAIPIIMSSVGCALILALITATLYKLGFPKRRHYKEMLEDKP 1140
Qy 1141 EDTATSFSGDDFSCVAPNVPLS 1161

Db 1141 EDTATSFSGDDFSCVAPNVPLS 1161
RESULT 2
US-09-891-943-2
; Sequence 2, Application US/09891943
; Publication No. US2003007278A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US2003007278A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/891,943
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1161
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-891-943-2
Query Match 100.0%; Score 5987; DB 10; Length 1161;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TFGTVLLSLVSLVHGNLNDVEEPTIFQEDAGFGQSVVQFGSRLVWGAFLVAAQNT 60
Db 1 TFGTVLLSLVSLVHGNLNDVEEPTIFQEDAGFGQSVVQFGSRLVWGAFLVAAQNT 60
Qy 61 GRLYDCAATGMCQIPPHIRPEAVNMSLGLTLASTNGSRLLACGPTLHRVCGENSYSK 120
Db 61 GRLYDCAATGMCQIPPHIRPEAVNMSLGLTLASTNGSRLLACGPTLHRVCGENSYSK 120
Qy 121 GSCLLLSRWELIQTVPDAPTECPHQMNDIVFLIDSGSIDQDNFNQMGFVQAVMGQFE 180
Db 121 GSCLLLSRWELIQTVPDAPTECPHQMNDIVFLIDSGSIDQDNFNQMGFVQAVMGQFE 180
Qy 181 GDTTLFALMOYNSLLKIHFTFTQFTSPSQSLVDPPIVOLKGLTPTATGILTVVTLFHH 240
Db 181 GDTTLFALMOYNSLLKIHFTFTQFTSPSQSLVDPPIVOLKGLTPTATGILTVVTLFHH 240
Qy 241 KNGARKSAKILIVITDGQKIDPLEYSDVIPQAEKAGIIRYAIGVGHAFQGPPTARQELN 300
Db 241 KNGARKSAKILIVITDGQKIDPLEYSDVIPQAEKAGIIRYAIGVGHAFQGPPTARQELN 300
Qy 301 TISSAPPQDHVKVDNFAALGSIQKQLEKIYAVGTSRASSSQHESQEGSFSTALTM 360
Db 301 TISSAPPQDHVKVDNFAALGSIQKQLEKIYAVGTSRASSSQHESQEGSFSTALTM 360
Qy 361 DGLFLGAVGSFWSGGAFLYPPNMSPTFINMSQENVMDRDSYLGYSTELALWKGVQNLVL 420
Db 361 DGLFLGAVGSFWSGGAFLYPPNMSPTFINMSQENVMDRDSYLGYSTELALWKGVQNLVL 420
Qy 421 GAPRYOHTGKAVITQVSRQWRKKAETVGTQIGSYFGASLCSVDVDSGSDTLILIGAPH 480
Db 421 GAPRYOHTGKAVITQVSRQWRKKAETVGTQIGSYFGASLCSVDVDSGSDTLILIGAPH 480
Qy 481 YTEQTRGGQVSVCPFRGQVQWQCDVLRGEGHPWGRFGAALTVLGDVNEKLDIVAI 540
Db 481 YTEQTRGGQVSVCPFRGQVQWQCDVLRGEGHPWGRFGAALTVLGDVNEKLDIVAI 540
Qy 541 GAPGEQENRGAVYLFHGAESGISPSHSQRIASSQLSPRLQYFGQALSGGQDLTQDGLMD 600
Db 541 GAPGEQENRGAVYLFHGAESGISPSHSQRIASSQLSPRLQYFGQALSGGQDLTQDGLMD 600

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2004, 09:37:59 ; Search time 23 seconds
(without alignments)
2605.989 Million cell updates/sec

Title: US-09-891-943-2
Perfect score: 5987
Sequence: 1 TFGTVLLSVLASYHGFLND.....DTATFGDDFSCVAPNVPLS 1161

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A COMB.pcp.*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pcp.*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pcp.*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pcp.*
5: /cgn2_6/ptodata/2/iaa/PCRUS COMB.pcp.*
6: /cgn2_6/ptodata/2/iaa/backfile1.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5987	100.0	1161	US-09-193-043-2	Sequence 2, Appli
2	5987	100.0	1161	US-09-688-307A-2	Sequence 2, Appli
3	5987	100.0	1161	US-09-350-259-2	Sequence 2, Appli
4	5981	99.9	1161	US-08-173-497-2	Sequence 2, Appli
5	5981	99.9	1161	US-08-286-889-2	Sequence 2, Appli
6	5981	99.9	1161	US-08-485-618-2	Sequence 2, Appli
7	5981	99.9	1161	US-08-362-652-2	Sequence 2, Appli
8	5981	99.9	1161	US-08-605-672-2	Sequence 2, Appli
9	5981	99.9	1161	US-08-482-293A-2	Sequence 2, Appli
10	5981	99.9	1161	US-08-943-363-2	Sequence 2, Appli
11	5971.5	99.7	1161	US-08-485-618-99	Sequence 99, Appl
12	5971.5	99.7	1161	US-08-605-672-99	Sequence 99, Appl
13	5971.5	99.7	1161	US-08-482-293A-99	Sequence 99, Appl
14	5971.5	99.7	1161	US-08-943-363-99	Sequence 99, Appl
15	5971.5	99.7	1161	US-09-193-043-99	Sequence 99, Appl
16	5971.5	99.7	1161	US-09-688-307A-99	Sequence 99, Appl
17	5971.5	99.7	1161	US-09-350-259-99	Sequence 99, Appl
18	4403.5	73.6	1161	US-09-193-043-55	Sequence 55, Appl
19	4403.5	73.6	1161	US-09-688-307A-55	Sequence 55, Appl
20	4403.5	73.6	1161	US-09-350-259-55	Sequence 55, Appl
21	4397.5	73.5	1161	US-08-485-618-55	Sequence 55, Appl
22	4397.5	73.5	1161	US-08-362-652-55	Sequence 55, Appl
23	4397.5	73.5	1161	US-08-605-672-55	Sequence 55, Appl
24	4397.5	73.5	1161	US-08-482-293A-55	Sequence 55, Appl
25	4397.5	73.5	1161	US-08-943-363-55	Sequence 55, Appl
26	4364.5	72.9	1151	US-08-286-889-37	Sequence 37, Appl
27	4364.5	72.9	1151	US-08-485-618-37	Sequence 37, Appl

28	4364.5	72.9	1151	1	US-08-362-652-37	Sequence 37, Appl
29	4364.5	72.9	1151	2	US-08-605-672-37	Sequence 37, Appl
30	4364.5	72.9	1151	2	US-08-482-293A-37	Sequence 37, Appl
31	4364.5	72.9	1151	2	US-08-943-363-37	Sequence 37, Appl
32	4364.5	72.9	1151	3	US-09-193-043-37	Sequence 37, Appl
33	4364.5	72.9	1151	3	US-09-688-307A-37	Sequence 37, Appl
34	4364.5	72.9	1151	4	US-09-350-259-37	Sequence 37, Appl
35	4364.5	72.9	1151	4	US-08-485-618-53	Sequence 53, Appl
36	4305.5	71.9	1161	1	US-08-362-652-53	Sequence 53, Appl
37	4305.5	71.9	1161	2	US-08-605-672-53	Sequence 53, Appl
38	4305.5	71.9	1161	2	US-08-482-293A-53	Sequence 53, Appl
39	4305.5	71.9	1161	2	US-08-943-363-53	Sequence 53, Appl
40	4305.5	71.9	1161	3	US-09-193-043-53	Sequence 53, Appl
41	4305.5	71.9	1161	4	US-09-688-307A-53	Sequence 53, Appl
42	4305.5	71.9	1161	4	US-09-350-259-53	Sequence 53, Appl
43	4278.5	71.5	1155	1	US-08-286-889-46	Sequence 46, Appl
44	4278.5	71.5	1155	1	US-08-485-618-46	Sequence 46, Appl
45	4278.5	71.5	1155	1	US-08-362-652-46	Sequence 46, Appl

ALIGNMENTS

RESULT 1
US-09-193-043-2
; Sequence 2, Application US/09193043
; Patent No. 6251395
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; TITLE OF INVENTION: No. 6251395el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/193,043
; CURRENT FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173,497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1161
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-193-043-2

Query Match 100.0%; Score 5987; DB 3; Length 1161;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	TFGTVLLSVLASYHGFLNDVEEPTIFQEDAGGFGQSVVQFGGSLVVGAPLEVVAANQT	60
Db	1	TFGTVLLSVLASYHGFLNDVEEPTIFQEDAGGFGQSVVQFGGSLVVGAPLEVVAANQT	60
QY	61	GRLYDCAATGMCQPIPLHPEAVNMSLGLTAASTNGSRLACGPTLHRVCGENSYSK	120
Db	61	GRLYDCAATGMCQPIPLHPEAVNMSLGLTAASTNGSRLACGPTLHRVCGENSYSK	120
QY	121	GSCLLGRSWEIIQVVPDAPTECPHOEMDIIVFLIDGSGSIDONDENOMKGFVQAVMGQFE	180
Db	121	GSCLLGRSWEIIQVVPDAPTECPHOEMDIIVFLIDGSGSIDONDENOMKGFVQAVMGQFE	180
QY	181	GTDTLFAIMOYNLKIHTFTQFRTSPSQSLVDPIVQLKGLTFTATGILTVVTLFPHH	240
Db	181	GTDTLFAIMOYNLKIHTFTQFRTSPSQSLVDPIVQLKGLTFTATGILTVVTLFPHH	240
QY	241	KNGARKSAKKILIVITDQKYPLEYSVDVPOAKAGIIRYATGVGHAFQGPTRQELN	300
Db	241	KNGARKSAKKILIVITDQKYPLEYSVDVPOAKAGIIRYATGVGHAFQGPTRQELN	300

301 TISSAPPDQHVFKVDNFAALGSIQKQOEKIYAVEGTQSRASSSFQHEMSQEGFSTALTM 360
Db TISSAPPDQHVFKVDNFAALGSIQKQOEKIYAVEGTQSRASSSFQHEMSQEGFSTALTM 360
361 DGLFLGAVGSGFSGGAFLLPPNMSPTFINNSQENVDMDSDYLGYSYTELALWKGVQNLVL 420
Db DGLFLGAVGSGFSGGAFLLPPNMSPTFINNSQENVDMDSDYLGYSYTELALWKGVQNLVL 420
421 GAPRYQHTGKAVIFTQVSRQWRKKAETGTQIGSYFGASLCSVDVSDGSDTLILIGAPH 480
Db GAPRYQHTGKAVIFTQVSRQWRKKAETGTQIGSYFGASLCSVDVSDGSDTLILIGAPH 480
481 YVEQTRGGQVSVCLPRGQVQWOCDAVLGEQGHGPHGRFGAALTVLGDVWNEKLDLVAI 540
Db YVEQTRGGQVSVCLPRGQVQWOCDAVLGEQGHGPHGRFGAALTVLGDVWNEKLDLVAI 540
541 GAPGEQENRGAVYLFHGAESGSGISPSHSQRIASSQLSPRLQYFGQALSGQDGLTQDGLMD 600
Db GAPGEQENRGAVYLFHGAESGSGISPSHSQRIASSQLSPRLQYFGQALSGQDGLTQDGLMD 600
601 LAVGARGQVLLRLSLPVLKVGAVMRSPVEVAVKVCWEKPSALBAGDATVCLTIQKS 660
Db LAVGARGQVLLRLSLPVLKVGAVMRSPVEVAVKVCWEKPSALBAGDATVCLTIQKS 660
661 SLDQGLDIQSSVRFDLALDPGLTSLRAIFNETKNPTLTRKTLGLGHCETLKLILPDCV 720
Db SLDQGLDIQSSVRFDLALDPGLTSLRAIFNETKNPTLTRKTLGLGHCETLKLILPDCV 720
721 EDVVSPIILHLNFSVLREPIPSQNLRLPVLAVGSDQLFTASLPFEKNCQDGLCEGDLGV 780
Db EDVVSPIILHLNFSVLREPIPSQNLRLPVLAVGSDQLFTASLPFEKNCQDGLCEGDLGV 780
781 TLFSPGSLQTLTVGSSLELNVITVWNADESGYGVVSLVYPAGLSHRRVSGAQKHQSA 840
Db TLFSPGSLQTLTVGSSLELNVITVWNADESGYGVVSLVYPAGLSHRRVSGAQKHQSA 840
841 LRLACETVPTDGLRSSRCVSNHPIFHEGSGNGTFIVFDVSKATLGRMLMRASASSE 900
Db LRLACETVPTDGLRSSRCVSNHPIFHEGSGNGTFIVFDVSKATLGRMLMRASASSE 900
901 NKAASSKATFQLELPVKYAVYTMISROBESTKYFNPAFSDKXKKEASHRYVNNLSOR 960
Db NKAASSKATFQLELPVKYAVYTMISROBESTKYFNPAFSDKXKKEASHRYVNNLSOR 960
961 DLAISINFWVPVLLNGVAVMDVMEAPSQLPCVSRKPPQHSDFLTQISRSPMLDCSTA 1020
Db DLAISINFWVPVLLNGVAVMDVMEAPSQLPCVSRKPPQHSDFLTQISRSPMLDCSTA 1020
1021 DCLQFRCDVPSFVQBELDFTLKNLSFGWVRETLQKVLVWSVAEITDTSVYSQLPQ 1080
Db DCLQFRCDVPSFVQBELDFTLKNLSFGWVRETLQKVLVWSVAEITDTSVYSQLPQ 1080
1081 EAFMRAQEMVLEDEYVNAIPIIMGSSVALLLALITATLYKLGFFKRYKEMLEKDP 1140
Db EAFMRAQEMVLEDEYVNAIPIIMGSSVALLLALITATLYKLGFFKRYKEMLEKDP 1140
1141 EDTATSGDDDFSCVAPNVPLS 1161
Db EDTATSGDDDFSCVAPNVPLS 1161

RESULT 2

US-09-688-307A-2
; Sequence 2, Application US/09688307A
; Patent No. 6432404
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 6432404el Human Beta-2
; FILE REFERENCE: 27866/36646
; CURRENT APPLICATION NUMBER: US/09/688,307A
; CURRENT FILING DATE: 2000-10-13

; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/605,672
; PRIOR FILING DATE: 1996-02-22
; PRIOR APPLICATION NUMBER: 08/173,497
; PRIOR FILING DATE: 1993-12-23
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1161
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-688-307A-2

Query Match 100.0%; Score 5987; DB 4; Length 1161;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TFGTVLLLSVLASYHGFLNDVEEPTIFQEDAGGFGQSVQFGGSRSLVVGAPLEVVAAQT 60
Db 1 TFGTVLLLSVLASYHGFLNDVEEPTIFQEDAGGFGQSVQFGGSRSLVVGAPLEVVAAQT 60
QY 61 GRLDYCAATACMCQOPIHLIRPEAVNMSLGLTLAASTNGSRLLACGPTLHRVCGENSYSK 120
Db 61 GRLDYCAATACMCQOPIHLIRPEAVNMSLGLTLAASTNGSRLLACGPTLHRVCGENSYSK 120
QY 121 GSCLLLSRWEEIIQTVPDATPECPHQEMDIVFLIDGSGSIDQDNFNQMKGFQVQVMOFE 180
Db 121 GSCLLLSRWEEIIQTVPDATPECPHQEMDIVFLIDGSGSIDQDNFNQMKGFQVQVMOFE 180
QY 181 GTDLFLALMQYNLLKIHFTFTQRTSPSQSLVDPIVOLKGLTFTATGILTVVTQLFHH 240
Db 181 GTDLFLALMQYNLLKIHFTFTQRTSPSQSLVDPIVOLKGLTFTATGILTVVTQLFHH 240
QY 241 KNGARKSAKLLIIVITDQKYKDPLEYSDDVIPOAEKAGIIRYAIYGVGHAFQGPTRAEIN 300
Db 241 KNGARKSAKLLIIVITDQKYKDPLEYSDDVIPOAEKAGIIRYAIYGVGHAFQGPTRAEIN 300
QY 301 TISSAPPDQHVFKVDNFAALGSIQKQOEKIYAVEGTQSRASSSFQHEMSQEGFSTALTM 360
Db 301 TISSAPPDQHVFKVDNFAALGSIQKQOEKIYAVEGTQSRASSSFQHEMSQEGFSTALTM 360
QY 361 DGLFLGAVGSGFSGGAFLLPPNMSPTFINNSQENVDMDSDYLGYSYTELALWKGVQNLVL 420
Db 361 DGLFLGAVGSGFSGGAFLLPPNMSPTFINNSQENVDMDSDYLGYSYTELALWKGVQNLVL 420
QY 421 GAPRYQHTGKAVIFTQVSRQWRKKAETGTQIGSYFGASLCSVDVSDGSDTLILIGAPH 480
Db 421 GAPRYQHTGKAVIFTQVSRQWRKKAETGTQIGSYFGASLCSVDVSDGSDTLILIGAPH 480
QY 481 YVEQTRGGQVSVCLPRGQVQWOCDAVLGEQGHGPHGRFGAALTVLGDVWNEKLDLVAI 540
Db 481 YVEQTRGGQVSVCLPRGQVQWOCDAVLGEQGHGPHGRFGAALTVLGDVWNEKLDLVAI 540
QY 541 GAPGEQENRGAVYLFHGAESGSGISPSHSQRIASSQLSPRLQYFGQALSGQDGLTQDGLMD 600
Db 541 GAPGEQENRGAVYLFHGAESGSGISPSHSQRIASSQLSPRLQYFGQALSGQDGLTQDGLMD 600
QY 601 LAVGARGQVLLRLSLPVLKVGAVMRSPVEVAVKVCWEKPSALBAGDATVCLTIQKS 660
Db 601 LAVGARGQVLLRLSLPVLKVGAVMRSPVEVAVKVCWEKPSALBAGDATVCLTIQKS 660
QY 661 SLDQGLDIQSSVRFDLALDPGLTSLRAIFNETKNPTLTRKTLGLGHCETLKLILPDCV 720
Db 661 SLDQGLDIQSSVRFDLALDPGLTSLRAIFNETKNPTLTRKTLGLGHCETLKLILPDCV 720
QY 721 EDVVSPIILHLNFSVLREPIPSQNLRLPVLAVGSDQLFTASLPFEKNCQDGLCEGDLGV 780